ECOLOGICAL MECHANISMS AND EFFECTIVENESS

OF BIOREMEDIATION IN ALASKA

By

Mary-Cathrine C. E. Leewis

RECOMMENDED:

Dr. Todd O'Hara

Dr. Roger Ruess

Dr. D. Lee Taylor

Dr. Mary Beth Leigh, Advisory Committee Chair

Dr. Diane Wagner, Chair
Department of Biology & Wildlife

APPROVED:

Dr. Paul Layer, Dean
College of Natural Science and Mathematics

Dr. John Eichelberger
Dean of the Graduate School

4/21/14
Date

ECOLOGICAL MECHANISMS AND EFFECTIVENESS OF BIOREMEDIATION IN ALASKA

A

DISSERTATION

Presented to the Faculty

of the University of Alaska Fairbanks

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

By

Mary-Cathrine C. E. Leewis, B.S., M.S.

Fairbanks, Alaska

May 2014

Abstract

What drives microbial community structure and function is a fundamental question of microbial ecology. Soil microbial communities have wide ranging metabolic capabilities, which include performing oxidation-reduction reactions responsible for cycling of nutrients and organic compounds and biodegradation of pollutants. One major determinant of microbial function in soils is vegetation type. Considering plants are diverse in chemical composition, they impact the quantity and quality of carbon and nutrients available to microbes through root turnover, root leachates, as well as by altering pH and soil microclimate (moisture, temperature). Rhizosphere interactions, in the form of phytoremediation, can be capitalized upon to provide a potentially cost effective method for detoxifying contaminated soils using plants and associated soil microorganisms. The remote locations and cold climate of Alaska provide unique challenges associated with phytoremediation such as finding effective plant species that can achieve successful site clean-up despite the extreme environmental conditions that includes minimal site management.

Here we investigate the potential mechanisms and related effectiveness of microbial communities and native boreal vegetation associated with contaminant degradation and biogeochemical cycling. We examined three different soil systems to understand how dominant vegetation type, historical treatment and contamination shape the microbial community structure and functional potential. First, we used stable isotope probing to understand how microbial communities act in concert to biotransform the recalcitrant contaminants, polychlorinated biphenyls. Second, we sought to understand if dominant vegetation type controls microbial

community structure and function either through direct impacts of plant root exudates and detritus or indirectly through the influence of plants on soil chemistry, composition, and structure. Finally, we conducted a forensic investigation of a petroleum contaminated site with no active site management for 15 years to assess the long-term effects of phytoremediation on soil petroleum concentrations, microbial community and vegetation colonization. The results of these experiments provide novel insights into the mechanisms of contaminant removal in boreal forest soils and the role of plants in ecosystem resilience to contamination, and demonstrates that phytoremediation using native and local plants can be an effective means to treat petroleum-contaminated soils.

Tables of Contents

Page

List of Figures

Page

List of Tables

Page

List of Appendices

Page

Acknowledgments

My research was generously supported by many funding agencies including: the University of Alaska Fairbanks Institute of Arctic Biology Summer Fellowship, University of Alaska Graduate School Thesis Completion Fellowship, Grant Number 5P20RR016466 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH) and graduate fellowships from the Alaska Idea Network for Biomedical Research Excellence (INBRE), and EPSCoR NSF award #EPS-0701898 & the State of Alaska. In addition, support for study at the Institute for Chemical Technology Prague was provided by the David L. Boren National Security Education Program (NSEP) Fellowship. Further support was provided by the Department of Biology and Wildlife and Department of Chemistry in the form of Teaching Assistantships and tuition waivers. I wish to thank my Advisory Committee chair, Dr. Mary Beth Leigh, for the intellectual and additional financial resources she provided throughout the duration of my project. Thoughtful comments from and discussion with Dr. Todd O'Hara, Dr. D. Lee Taylor, and Dr. Roger Ruess offered valuable perspective on the complex data set collected for my research, and I thank them all for their help.

I thank Ondřej Uhlík, Tomaš Macek, Martina Macková, and the entire lab at VSCHT, Prague for their helpful discussions and cultural exchange. I would like to thank Jakub Rídl and Miluse Hroudová of the Institute of Molecular Genetics, Czech Academy of Sciences for help with microbial community sequencing for both Chapters Four and Five. I also thank Dr. Joy VanNostrand and the Institute for Environmental Genomics at the University of Oklahoma for help with the GeoChip sample preparation and data processing associated with Chapter Five.

Chapter Six would not have been possible without the work done by Dr. C. Mike Reynolds and the Army Corps of Engineers Cold Regions Research and Engineering Laboratory. I also wish to thank Chris Kasanke and Carl Richmond for help with field work and sample processing and Jan Fišer, Iva Pacovská from the Institute for Chemical Technology Prague for help with lab work associated with Chapter Six. Earlier site and laboratory work for Chapter Six were supported by Army Environmental Quality Technology Program and the Alaska Science and Technology Foundation. In addition, I would like to thank Jacobs Engineering for access to soil samples used in Chapter Four.

Finally, I would like to thank my family for their loving and unconditional support over the years; my father and mother, Keith Leewis and Mary-Jane Leewis-Kirk, my sister and her family, Alice, Baylee and Ian (&newbie) Nicholls, my brothers' family, Owen Leewis and Jennifer Campbell, and my faithful companion Fred.

Chapter 1

General Introduction

With the ongoing global rise in temperature, there is increased scientific and economic interest in arctic and boreal regions (1). Expanding development and transportation activities are raising the risk of disturbance from accidental spills of petroleum hydrocarbons used for heating, transportation fuels, and electricity, and from increased natural resource exploration, such as crude oil, coal and natural gas (2). The region is already home to hundreds of legacy contaminated sites from previous military and development activities (3). There are more than 500 formerly used defense sites owned by the Department of Defense in Alaska with approximately 1000 individual areas of soil contamination, many contaminated with petroleum (3, 4) as well as polychlorinated biphenyls (PCBs) and other contaminants. Petroleum spills are also widespread throughout the state and region as a result of fuel transport, storage, and use.

Soil microbial communities have wide ranging metabolic capabilities, which include performing oxidation-reduction reactions responsible for cycling of nutrients (nitrogen, phosphorus, sulfur) and organic compounds (lignocellulose) that can include biodegradation of pollutants (5). A powerful determinant of microbial function in soils is vegetation type (6, 7). Because plants are diverse in their chemical composition, they impact the quantity and quality of carbon and nutrients available to microbes through root turnover, root leachates, as well as by altering pH and soil microclimate (moisture, temperature) (8). Understanding the relationship between vegetation type and microbial community function is important to understanding

ecosystem function and understanding how it may respond to large scale disturbances such as a contamination events and rising temperatures.

This dissertation includes three research-based chapters, and two appendices. In Chapter 2, I investigated carbon flow through a microbial community during the biodegradation of some polychlorinated biphenyls (PCB) within a contaminated soil microbial ecosystem from Kodiak Island, Alaska. PCBs are toxic persistent organic pollutants (POPs) that some soil microorganisms have been shown to degrade, however there is little current understanding of these interactions with the microbial community active in different steps in PCB biodegradation. DNA-SIP was used to assess the bacterial community involved in degradation of ^{13}C-biphenyl (BP), which is an analogue of PCBs, and ^{13}C-benzoate (BZ), a product/intermediate of biphenyl degradation and analogue of chlorobenzoates. Using SIP performed in parallel with multiple substrates over time, the study reveals microbes performing the upper (BP) and lower (BZ) degradation pathways, as well as heterotrophic bacteria involved indirectly in biodegradation. Biotransformation of BP and BZ in soils was mainly mediated by *Actinobacteria* and *Proteobacteria*. *Acidothermus cellulolyticus, Gemmatimonas aurantiaca*, and *Pseudomonas lutea* were active in processing carbon derived from BP, while *Burkholderia* spp., *Methylobacterium jeotgali,* and *Rhodococcus qingshengii* were dominant in processing BZ-derived carbon. In addition, there were also genera detected which derived carbon from both labeled BP and BZ including *Arthrobacter*, *Methylobacterium, Afipia, Gemmatimonas,* and *Rhodanobacter.* The microbial populations found to degrade BZ had lower diversity and utilized more substrate than the BP-utilizing populations. Overall, the data show that biotransformation of BP and BZ was mainly performed by microorganisms with apparently different growth

strategies: BZ-associated bacteria are fast growing, copiotrophic organisms, while microbes which transform BP are oliotrophic, slower growing, organisms.

The third chapter investigates the microbial community structure and functional potential in tree stands dominated by four different vegetation types within the Alaskan boreal forest. Microbial DNA was extracted from soils collected from the Bonanza Creek Long Term Ecological Research Site (BNZ LTER) in interior Alaska. Using a combination of functional gene microarray and next generation sequencing techniques, I screened the microbial structural and functional potential of root zone soils associated with four vegetation types dominated by either *Picea glauca* (white spruce), *Picea mariana* (black spruce), *Populus tremuloides* (quaking aspen), or *Betula neoalaskana* (Alaskan paper birch), located within the Bonanza Creek Long Term Ecological Research Site (BNZ LTER) in interior Alaska. Results indicate that microbes with the genetic functional potential to degrade chlorinated solvents, hydrocarbons, herbicides, pesticides, and other aromatic compounds are present in varying amounts in the root zone of each tree species, despite being located in pristine forests. Results also indicated that microbial communities and functional metagenomes vary in composition in accordance with the dominant tree species present. This study provides insight into the role of plants in driving microbial community structure and functional potential, and suggests that substantial resilience to contaminants and phytoremediation potential exists in Alaskan boreal forest soils.

The final research chapter describes the long term effects of phytoremediation on petroleum contaminated soils in Fairbanks, Alaska. A phytoremediation field study was initiated in 1995 and 1996 at the Farmers Loop Permafrost Research Facility field site of the Army Corps of Engineers Cold Regions Research and Engineering Laboratory (ACE CRREL) in Fairbanks,

Alaska. The study sought to compare the effects of nutrients and cold-hardy grasses on the microbial biodegradation of petroleum contaminated soils. After the initial 2-year study, the site was no longer actively managed and a mixture of local native, naturalized and invasive plants colonized the plots. The field site was re-examined after 15 years with no active site management to assess the long-term effects of phytoremediation on a) colonization by native and non-native plants b) their effects on soil bacteria and c) petroleum removal from soil. Native and non-native vegetation had extensively colonized the site and none of the originally planted grass species were present on the site after 15 years. Total petroleum hydrocarbon (TPH) concentrations achieved regulatory clean up levels in both soil types and all treatment groups, with lower TPH concentrations correlating with higher amounts of woody vegetation (trees and shrubs). The original treatment type affected vegetation recruitment to each plot with woody vegetation and more native plants in unfertilized plots. Bacterial community structure also varied according to the originally applied treatments. The results indicate that the addition of plants and/or fertilizer sent the plots on different plant successional trajectories, resulting in different plant community assemblages and associated microbial communities. Regardless of treatment, the petroleum was remediated to below regulatory limits within 15 years, including in initially unplanted plots. The initial 2-3 year study demonstrated that dense grass growth accelerated contaminant biodegradation, but the grasses used did not prove to be persistent at the site. The positive correlation between woody vegetation and petroleum loss in later years suggests that trees and shrubs may be more effective for accelerating petroleum biodegradation. This study suggests that initial treatment with native tree species in combination with perennial grasses could be an effective means for phytoremediating petroleum-contaminated soils and promoting ecological recovery.

The first appendix is a review of multiple molecular approaches which can be used for the analysis of microbial, especially bacterial, diversity in contaminated environments. Until recently, knowledge of the microbial communities involved in bioremediation processes had been based mostly on culture-dependent studies, which do not account for the fact that laboratory conditions differ substantially from the environment. Cultivation-based techniques have been found to detect only about 1% of microbes occurring in the environment (9). Over the last two decades, the development of molecular techniques has enabled culture-independent phylogenetic analyses of microbial communities and functional genes. These approaches have demonstrated that bacterial diversity is far more extensive than determined solely based on culture-based studies.

The second appendix is a detailed review of the molecular technique, stable isotope probing (SIP), in the context of contaminant degradation. Although the development of culture-independent techniques for analysis of microbial communities has allowed for more detailed analysis of microbial communities, the linkage of contaminant transformation to phylogenetic identity and specific genes/enzymes of metabolically active microbes without cultivation remained a major challenge and still required cultivation. The development of SIP was instrumental in circumventing the limitations of culture based investigations of substrate biotransformation. SIP enables the direct linkage of microbial metabolic capacity to phylogenetic and metagenomic information within a community context by tracking isotopically labeled substances into phylogenetically and functionally informative biomarkers.

The primary objectives of this dissertation are to a) review methods for assessing microbial ecology of contaminated sites, b) identify PCB-degrading microorganisms and

characterize how members of the microbial community interact to biodegrade PCBs in contaminated soil, c) characterize the microbial community structure and functional potential associated with several vegetation types within the Alaskan boreal forest, including contaminant biodegradation and carbon cycling, and d) determine long-term effectiveness of bioremediation/phytoremediation on petroleum contaminated soils at in interior Alaska. Together, the results provide novel insights into the mechanisms of contaminant removal in boreal forest soils, the role of plants in ecosystem resilience to contamination, and demonstrates that phytoremediation using native and local plants can be an effective means to treat petroleum-contaminated soils.

REFERENCES

1. **Yergeau E, Sanschagrin S, Beaumier D, Greer CW**. 2012. Metagenomic analysis of the bioremediation of diesel-contaminated Canadian high arctic soils. PLoS One **7**:e30058.

2. **Leewis M-C, Reynolds CM, Leigh MB**. 2013. Long-term effects of nutrient addition and phytoremediation on diesel and crude oil contaminated soils in subarctic Alaska. Cold Reg. Sci. Technol. **96**:129–137.

3. **Reynolds CM, Koenen BA**. 1997. Rhizosphere-Enhanced Bioremediation. Mil. Eng. **586**:32–33.

4. **Yergeau E, Lawrence JR, Sanschagrin S, Waiser MJ, Korber DR, Greer CW**. 2012. Next-generation sequencing of microbial communities in the Athabasca River and its tributaries in relation to oil sands mining activities. Appl. Environ. Microbiol. **78**:7626–7637.

5. **Nannipieri P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G**. 2003. Microbial diversity and soil functions. Eur. J. Soil Sci. **54**:655–670.

6. **Thomson BC, Ostle N, McNamara N, Bailey MJ, Whiteley AS, Griffiths RI**. 2010. Vegetation affects the relative abundances of dominant soil bacterial taxa and soil respiration rates in an upland grassland soil. Microb. Ecol. **59**:335–43.

7. **Hartmann A, Schmid M, Tuinen D Van, Berg G**. 2008. Plant-driven selection of microbes. Plant Soil **321**:235–257.

8. **Leigh MB, Fletcher JS, Fu X, Schmitz FJ**. 2002. Root turnover: an important source of microbial substrates in rhizosphere remediation of recalcitrant contaminants. Environ. Sci. Technol. **36**:1579–83.

9. **Torsvik V, Øvreås L**. 2002. Microbial diversity and function in soil: from genes to ecosystems. Curr. Opin. Microbiol. **5**:240–5.

Chapter 2

Aromatic Compound Processing in PCB-Contaminated Soil: Carbon Flow Through the Microbial Ecosystem [1]

ABSTRACT

Polychlorinated biphenyls (PCBs) are toxic and persistent organic pollutants that some microorganisms can biotransform in soil systems. However, there is little current understanding of the interactions between members of the microbial community active in PCB biodegradation. This study investigates microbial biodegradation and carbon flow through PCB-contaminated soils from Kodiak Island, Alaska. DNA-stable isotope probing (SIP) was used to identify the bacterial community involved in degradation of ^{13}C-biphenyl (BP, an analogue of PCBs), and ^{13}C-benzoate (BZ a product/intermediate of biphenyl degradation). By using SIP in parallel with multiple substrates and over a time course, we reveal microbes performing the upper (BP) and lower (BZ) degradation pathways, as well as heterotrophic bacteria involved indirectly in biodegradation. Biotransformation of BP and BZ in soils was mainly mediated by *Actinobacteria* and *Proteobacteria*. *Acidothermus cellulolyticus, Gemmatimonas aurantiaca,* and *Pseudomonas lutea* were active in processing carbon derived from BP, while *Burkholderia* spp., *Methylobacterium jeotgali,* and *Rhodococcus qingshengii* were dominant in processing BZ-derived carbon. The microbial populations found to degrade BZ had lower diversity and utilized

[1] Leewis, M.C., O. Uhlik, M.B. Leigh. Aromatic Compound Processing in PCB-Contaminated Soil: Carbon Flow Through the Microbial Ecosystem. Prepared for submission to Applied and Environmental Microbiology.

more substrate carbon than the BP-utilizing populations. Overall, our data suggest that biotransformation of BP and BZ was mainly performed by microorganisms with different growth strategies: BZ-associated bacteria are fast growing, copiotrophic organisms, while microbes which transform BP are oliotrophic, slower growing, organisms.

INTRODUCTION

Polychlorinated biphenyls (PCBs) are a class of persistent organic pollutants (POPs) that were widely used in applications such as heat transfer fluids, electrical insulators and lubricating fluids until their production was banned globally in the 1980s. It is estimated that almost ten million kilograms still persist in the environment (1). Widespread contamination due to spills, improper disposal and abandoned machinery persist in the environment creating a threat to human health and the environment. In Alaska, high-level PCB contamination has been found at over 35 formerly used defense sites (FUDS) (2). Two potential remediation strategies for PCB contaminated sites are *ex situ* physical-chemical remediation by contaminant removal and bioremediation, which can be performed *in situ*, through the use of microorganisms and plants (3). Cleanup of sites through physical removal of soils and transport to a treatment facility can be extremely costly and result in much of the contaminant left behind in association with loose soil (4). Biological methods of contaminant removal, such as microbial bioremediation, are a potentially less expensive alternative for soil cleanup (5–7).

For more than three decades, the aerobic microbial degradation of PCBs, its analogue biphenyl (BP) (8) and its degradation intermediate benzoate (BZ) has been the subject of a large body of research (9). The aerobic biodegradation of (chloro)biphenyl to (chloro)benzoate and (chloro)2-hydroxypenta-2,4-dienote, termed the upper PCB degradation pathway, is characterized by four separate reactions with the initial reaction and attack of a BP 2,3-dioxygenase (*bphAE* or *bphA1A2A3A4*) (3, 9, 10). Enzymatic attack of chlorinated BP is dependent on the substitution pattern, with congeners with fewer chlorine substitutions generally being more readily degraded than BP with more chlorine substitution and BP with chlorine on one aromatic is ring easier to degrade than those with chlorines on both rings (9). The lower

pathway of PCB degradation begins with the transformation of 2-hydroxypenta-2,4-dienoate by 2-hydroxypenta-2,4-dienoate hydratase (*bphH*), 4-hydroxy-2-oxovalterate aldolase (*bphJ*) and an acylating acetaldehyde dehydrogenase (*bphI*) to acetyl-CoA, which can then enter the Krebs cycle and (chloro)BZ (9). (Chloro)BZ is further metabolized via catechol and the β–ketoadipate pathway and funneled into central metabolism (9, 11, 12). Some microorganisms, such as certain *Arthrobacter, Pseudomonas, Rhodococcus* and *Burkholderia*, are capable of performing complete mineralization of BP and some chlorobiphenyls (13, 14) and growing on a range of chlorinated substrates such as chlorinated BZ and BP (14). Other organisms are only known to possess the upper pathway and/or to generate dead-end metabolites of BP (9, 15). BZ catabolic pathways such as the β-ketoadipate pathway are relatively widespread in the soil environment (12). In the context of the contaminated soil environment, it is not known if BP is processed synergistically by separate microbial guilds specializing in BP and BZ biodegradation.

Understanding the mechanisms underlying microbial community processing of complex carbon substrates has long been a goal of microbial ecologists, however this task is challenging for a variety of reasons. Until recently, it has been difficult to link bacterial community taxonomy to function due to the limitations associated with isolating and cultivating most microbes in the laboratory. Stable isotope probing (SIP) methods (16–18) have greatly advanced understanding which organisms within a complex community are active in degrading certain substrates, whether directly or indirectly. SIP enables the tracking of isotopically labeled C, N or O into organisms that perform primary biodegradation of the substrate, as well as those that benefit from secondary feeding on intermediates or biomass of primary degraders. Techniques which isolate isotopically-labeled biomarkers allow for the downstream analysis of microorganisms that have incorporated the substrate, which implicates them in the degradation

of the compound or its metabolites. SIP, as a culture independent method, is an effective tool for linking microbial community structure and function. To date, SIP studies have provided valuable insight into the processing of C from individual substrates or mixtures of compounds (e.g. plant litter) (11, 19–21). SIP also has great unrealized potential to reveal how the microbial community interacts with complex carbon substrates to degrade the compound synergistically and to show how carbon from a complex substrate flows through different microbial guilds over time (22, 23).

In order to understand natural biodegradative processes in contaminated environments, it is important to identify the microorganisms involved and elucidate their growth strategies. This study aims to reveal information concerning synergistic biodegradation and carbon flow in weathered PCB contaminated soil from Kodiak, AK. In this study, we employed DNA SIP in parallel incubations with two different substrates to identify the microbes performing the upper (BP) and lower (BZ) PCB biodegradation pathways. We then used bacterial, 16S rRNA gene, community pyrosequencing to identify the individual bacterial taxa that derived carbon from the substrates over time, including those that actively processed carbon from both substrates. We hypothesized that since not all bacteria possess genes for the upper and lower pathways; (1) a portion of the bacterial community will be active only in either the upper or lower pathway of BP degradation, however (2) a subset of the community will be able to complete the entire degradation pathway, and (3) given the widespread nature of BZ degradation genes, a larger proportion of the community will be involved in BZ degradation than BP degradation. To our knowledge, this work represents the first study to actively track carbon flow through a PCB contaminated system in a sub-Arctic environment.

MATERIALS AND METHODS

Soil Sample. Soil was collected from the Coast Guard Station at Drury Gulch, Kodiak Island, Alaska, USA in 1999 as a portion of clean up actions in accordance with Federal Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA) as amended by the Superfund Amendments and Reauthorization Act of 1986 (SARA) (http://dec.alaska.gov/spar/csp/dod_sites.htm). Soils were collected in clean sterile glass jars and stored at 4°C in glass jars until analysis in 2008. On average, PCB concentrations at the site ranged from 6.8 mg/kg to 14.0 mg/kg, with at least one hot spot of soil with 103,000 mg/kg PCBs. Cadmium, chromium, and lead were also identified at the site at concentrations of 5.0 mg/kg, 40.8 mg/kg and 541 mg/kg, respectively (24).

PCB Extraction and Quantification. The Soxhlet method was used to extract PCBs from aliquots of soils used in the molecular analyses according to the method of Slater et al. 2011 (25). Calculated PCB concentrations in investigated soils was 32.53 mg/kg.

Stable Isotope Probing (SIP) Microcosm Setup. Microcosms for SIP incubations were established in sterile 125 mL glass serum bottles (Wheaton, Millville, NJ, USA). The ^{13}C BP, ^{13}C BZ (Isotec, Miamisburg, OH, USA), unlabeled BP, or unlabeled BZ were dissolved in acetone and an appropriate volume was added to each microcosm by pipetting the solution onto the inside of the serum bottle and allowing the acetone to evaporate, leaving behind 1 mg BP or BZ crystals on the walls of the serum bottles (19). Soils were homogenized and added to microcosms after BP or BZ solutions had dried fully. Added to each serum bottle were 7 g of PCB contaminated soil from Drury Gulch (Kodiak Island, Alaska, USA) and 400 µL sterile water. The bottles were then sealed with a Teflon stopper and aluminum crimp top. Duplicate microcosms

were set up for each substrate (labeled or unlabeled BP or BZ) and each incubation time point (1, 4 and 14 days). These time points were selected based on times previously shown to reveal BP degrading bacteria using SIP (19). Microcosms were then incubated at room temperature (approximately 22°C) until destructive harvesting. An aliquot of soil without the addition of ^{13}C labeled substrate was stored at 4°C to serve as the time 0 (T0) sample for comparison with incubated samples. At the time of harvest for each microcosm, a 1 mL headspace gas sample was collected through the Teflon stopper with a sterile needle and syringe, which was then injected into an airtight gas sampling tube (12 mL) equilibrated with ambient lab air at the same time microcosms were sealed. These gas samples were then submitted for isotopic analysis of CO_2 as described below. Microcosms were harvested at 1, 4, and 14 days (d01, d04, and d14 respectively) and soil was stored at -80°C until further analysis.

Headspace Gas Isotope Analysis. The $\delta^{13}C$ was determined for headspace CO_2 gas samples at the Alaska Stable Isotope Facility (University of Alaska Fairbanks) using a Thermo Finnigan GasBench II carbonate analyzer with a DeltaPlusXP Mass Spectrometer. Reactions were run at 70°C with a He flow rate of 120 mL/min and a GC oven temp of 50°C. Standard sample tubes were purged for 20 minutes with UHP Helium and then 0.2 mL of 85% H_3PO_4 was manually added to the standards by syringe to the sample through the septum in the lid of the tubes. CO_2 gas samples were analyzed directly, with no headspace flushing or acid addition. Headspace gases were then transferred to the GasBench II where the water was removed through a nafion dryer. CO_2 was separated chromatographically from the other gases present and then transferred to the isotope ratio mass spectrometry (IRMS), where carbon were measured. $\delta^{13}C$PDB values are reported in reference to international isotope standards. Quality control

measures were taken by analyzing a blank every twenty samples. The laboratory working standard was calcium carbonate.

DNA Extraction. Total soil DNA was extracted using the PowerMax soil DNA isolation kit (MoBio Laboratories, Inc., USA) per instructions with the following modifications adapted from Uhlik et al. (2009) to increase DNA yield. After the final elution the DNA was concentrated by adding 0.2 mL 5M NaCl and 10.4 mL ethanol, incubated at -20°C overnight. After incubation the solution was gradually transferred into 2 mL microcentrifuge tubes with 20 μg glycogen (Roche, Germany), which were centrifuged after each addition in order to obtain a single pellet. The pellet was then dissolved into 50 μL sterile water. DNA concentrations were evaluated by measuring absorbance at 260 and 280 nm using a NanoDrop ND-1000 Spectrophotometer (Thermo Fischer Scientific, USA). All solutions were diluted to obtain a concentration of 100 ng/μl.

Isopycnic Centrifugation and Gradient Fractionation. Five microliters of each DNA extract (100 ng/μl) was mixed with 3.3 mL cesium trifluoroacetate (Amersham, United Kingdom) with a density of 1.6 g/ml and subjected to isopycnic density gradient centrifugation. Density gradient centrifugation was performed on the DNA extracted from each time point (T0-T14) in order to separate the ^{13}C-DNA from microbes that utilized the labeled carbon source from the rest of the soil community's unlabeled DNA. Isopycnic centrifugation, fractionation and precipitation following methods previously described (26). Fractions were frozen at -20°C until further analysis.

Quantitative PCR (q-PCR) screening of SIP Fractions. To locate and quantify bacterial DNA in SIP fractions, q-PCR was performed on every fraction for each time point in

duplicate as described previously (26, 27). Universal bacterial primers targeting conserved regions of eubacteral 16S rRNA 786f (5'-GATTAGATACCCTGGTAG-3') and 939r (5'-CTTGTGCGGGCCCCCGTCAATTC-3') were used for q-PCR. Gene copy number was determined using a standard curve constructed with *Pseudomonas stutzeri* JM300 genomic DNA as described previously (Uhlik et al., 2009). Fractions that were determined to contain ^{13}C-labeled bacterial DNA based on quantities of DNA in high density fractions that exceeded that of unlabeled controls were combined to create a "heavy" fraction (BD 1.584 – 1.628 $g.ml^{-1}$) and analyzed.

T-RFLP Analysis. Bacterial community profiling using T-RFLP was performed on heavy fractions (^{13}C-DNA), control heavy fractions (^{12}C incubated soils), and total community DNA to provide an initial overview of the diversity of labeled DNA and to assess any background contamination in heavy fractions from unlabeled DNA prior to pyrosequencing. PCR reactions were carried out as described previously (20, 27). Terminal restriction fragments less than 50 and above 1000 bp were eliminated from all data sets.

Phylogenetic identification of bacteria using bar-coded pyrosequencing. Amplification and pyrosequencing were performed by a modification of a method described previously by Engelbrektson et al. (2010) (28). A portion of the 16S rRNA gene (positions 8 to 357 [V1 and V2], *Escherichia coli* numbering) was amplified using the 8F primer containing a Roche 454 A pyrosequencing adaptor (underlined) and a 10 base pair unique bar code (see supplemental materials) and 357R primer also containing a Roche 454 B pyrosequencing adaptor (underlined): 8F 5'- CGTATCGCCTCCCTCGCGCCATCAG-NNNNNNNNNN-AGAGTTTGATCMTGGCTCAG -3', 357R 5'-

CTATGCGCCTTGCCAGCCCGCTCAGCTGCTGCSYCCCGTAG -3'. The MID (bar code) for each sample was unique to facilitate sorting of sequences from a single pyrosequencing run.

The PCR mixture (final volume, 25 µl) contained 1 µl each primer (10 µM), 0.5 µl dNTP mix (10 mM), 2.5 µl FastStart 10 X Buffer #2, 0.25 µl FastStart HiFi Polyumerase (5 U/µl), 18.75 µl molecular biology grade water (Roche). The following thermal cycling scheme was used: initial denaturation at 95°C for 3 min and 30 cycles of denaturation at 95°C for 30 sec, annealing for 1 min at 55°C, and extension at 72°C for 1.5 min, followed by a final extension period at 72°C for 10 min. Each PCR product was obtained in three parallel reactions, the resulting preparations were mixed, purified using Pure-Link PCR purification kit (Invitrogen, USA) and pooled for downstream sequencing. Roche 454 GS FLX Titanium sequencing (454 Life Sciences, USA) was performed on pooled reactions at the Institute of Molecular Genetics, Czech Academy of Sciences, Prague, Czech Republic.

Processing of Pyrosequencing Data. Raw pyrosequencing data (*.sff files) were processed using the mothur software package version 1.30.2 (29) and following the standardized operating procedure (29) with minor modifications (11). In brief, sequences were discarded if there was more than one mismatch in the bar code, more than two mismatches in the primer sequence, more than eight homopolymers, or were less than 200 bp. Sequence control and analyses were performed on 16S rDNA from ^{13}C-DNA from the BP microcosms day 4 and 14, BZ day 1 and 4 according to He et al. (20, 26, 27). The representative sequences have been deposited to the NCBI Short Read Archive under the accession number TBA.

RESULTS

Excess $^{13}CO_2$ was detected in microcosm headspaces within 1 day of incubation and increased in abundance throughout the experiment (14 days). No change in $^{13}CO_2$ was detected in control microcosms containing unlabeled BP or BZ. Based on previous reports that bacterial cells growing on BP mineralize 60-80% of BP to CO_2 (30), we calculated the amount of ^{13}C-BP or ^{13}C-BZ utilized by the microbial community (Table 2-1). In the ^{13}C-BP incubated soils, 0.24-0.32 µg of ^{13}C was metabolized within one day, and up to 0.79-1.05 µg within 14 days. In ^{13}C-BZ exposed populations, the amount of ^{13}C metabolized was 0.54-0.72 µg within one day and was 12.5-16.7 µg within a 14-day incubation.

Q-PCR analyses of gradient fractions were conducted to examine the relative abundance of 16S rRNA gene throughout the gradient and to identify fractions containing ^{13}C-labeled DNA (Figure S1). There was a substantial peak of DNA at buoyant density (BD) of 1.576 g.ml^{-1}, which corresponds to the BD of unlabeled (predominantly ^{12}C) DNA. In ^{13}C-incubated samples, there were two peaks detected indicating the presence of ^{13}C -DNA. A small amount of bacterial DNA was detected in fractions with BD higher than 1.584 g.ml^{-1} for unlabeled control DNA. Based on the q-PCR results, we combined the fractions with BD of 1.584-1.628 g. ml^{-1} as compiled heavy fractions prior to downstream molecular analyses as density gradients run in duplicate for each sample were similar in DNA quantity and distribution (Figure S2-1).

T-RFLP analyses of the ^{13}C-DNA from incubated soils indicated that a diverse array of bacteria derived carbon from ^{13}C-BZ or ^{13}C-BP from all time points and that many peaks were present in labeled DNA from both substrates. Only a small subset of peaks were found in the heavy fractions from control DNA, and fragments of the same size were not major contributors to T-RFLPs in the heavy fractions from ^{13}C-incubated samples (Figure S2-2). This indicated that

background DNA contamination in heavy fractions was minimal. As incubation time increased, each of the ^{13}C-DNA samples displayed fewer peaks with higher magnitudes (Figure S2-2).

To identify the bacterial populations active in assimilating carbon from labeled substrates, we sequenced the 16S rRNA gene amplicon (regions V1 and V2) using pyrosequencing. The numbers of pyrosequencing reads of labeled DNA in the samples selected ranged from 5684 to 14848 per sample after sequence processing (Table 2-2). Rarefaction curves indicated that OTU richness was saturated in the ^{13}C-BZ incubated samples. In the ^{13}C-BP incubated samples, however, the rarefaction curves do not reach an asymptote, and therefore the total diversity has not been fully sampled (Figure S2-3). The rarefaction curves also indicated that full community coverage was not achieved in the total soil community sequence library, however this is to be expected because soil communities are highly diverse.

Phylogenetic affiliations of pyrosequencing reads from ^{13}C-DNA by RDP classifier showed that diverse bacterial populations were involved directly or indirectly in acquiring BP or BZ-derived carbon (Table S2-1). Sequences from the phylum *Actinobacteria* predominated in the total community and ^{13}C incubated BZ day 4, and BP days 4 and 14 communities. Sequences of the phylum *Proteobacteria* were present in all of the examined time points, but predominated in the BZ day 01 sample, where they accounted for 36.5% of total pyrosequencing reads. In addition, the phyla *Acidobacteria, Actinobacteria, Gemmatimonadetes,* and *Nitrospira* were found in all of the examined time points and ^{13}C-DNA. In the early (d01) ^{13}C-BZ treated soils, sequences of the order *Rhizobiales* predominated the labeled DNA, whereas *Burkohlderia* and *Actinomycetales* dominated after 4 days of ^{13}C-BZ. In the early (day 04) ^{13}C-BP treated soils, *Acidobacteria* predominated labeled DNA and after 14 days, labeled sequences from ^{13}C-BP soils were dominated by *Actinobacteria*. Some pyrosequencing reads could not be assigned by

the RDP classifier, with 16.6% of the total community, 5.1-6.0% of the ^{13}C-BZ, and 13.2-19.9% of ^{13}C-BP pyrosequencing reads remaining unclassified.

The diversity of bacteria that derived C from ^{13}C-BP, as measured by Chao1, Inverse Simpson and effective number of species, decreased throughout the incubation. Effective number of species was calculated by taking the exponent of the Shannon index, according to Jost (2006, 31). Samples incubated with ^{13}C-BP had the highest diversity after four days of incubation. Bacteria labeled through incubation with ^{13}C-BZ also decreased in diversity over time and, bacterial sequences associated with the four day incubation had the lowest diversity of all measured populations (Table 2-2).

The percentage of unique and shared OTUs (defined at 97% sequence similarity) between communities that derived carbon from the different substrates and time points is represented using Venn diagrams in Figure 2-1. When all OTUs present in ^{13}C-DNA from BP and BZ were considered, 16.4% of the OTUs were shared in labeled DNA from both substrates. In soils incubated with ^{13}C-BP, 24.3% of OTUs were shared between the early and late incubated ^{13}C-BP soils, and 60.9% of unique OTUs were present only in the early (day 4) incubated soils, while the later harvested (day 14) soils had only 14.8% unique OTUs. In ^{13}C-BZ incubated soils, 15.5% of sequences were shared among early (day 1) and late (day 4) time points, and the early harvested soils (day 1) had fewer (27.9%) unique OTUs than the late harvested (day 4) soils with 56.6% unique OTUs. To examine the potential transition between the upper and lower portions of the BP degradation pathway, we compared percentage of unique OTUs in the late (day 14) ^{13}C-BP and early ^{13}C-BZ (day 1) incubated soils: 12.6% of OTUs were shared between the two pathways.

The abundance and relatedness of the top 200 taxa present in all of the samples were measured (Figure 2-2). The two communities incubated with ^{13}C-BP were more similar to each other than to communities incubated with ^{13}C-BZ. Taxa from the early incubated (day 1) ^{13}C-BZ were more similar to both the ^{13}C-BP incubated soils, and the later ^{13}C-BZ incubated soils are the most unique of the incubated samples. The data also indicated that there were fewer labeled taxa in the ^{13}C-BZ incubated soils than those incubated with ^{13}C-BP.

To investigate the phylogenetic differences and overlaps between BZ and BP utilizers, we generated a table for the top ten most abundant OTUs from each incubated time point (Table 2-3). The dominant OTUs labeled with ^{13}C-BP after four days (early) of incubation clustered with *Acidothermus cellulolyticus* (OTU 5, 4%), *Gemmatimonas aurantiaca* (OTU 9, 2%), *Schlegelella thermodepolymerans* (OTU 14, 2%), *Nocardioides jensenii* (OTU 13, 2%), and *Aciditerrimonas ferrireducens* (OTU 19, 2%). In soils incubated with ^{13}C-BP after fourteen days (late) of incubation the dominant taxa clustered with *Arthrobacter oryzae* (OTU 2, 12%), *Thermosipho geolei* (OTU 7, 5%), *Pseudomonas lutea* (OTU 24, 4%), *Acidothermus cellulolyticus* (OTU 5, 3%), and *Rhodanobacter soli* (OTU 17, 2%). Sequences clustering with *Acidothermus cellulolyticus* (OTU 5) were present in appreciable amounts in both the early (day 4) and late (day 14) ^{13}C-BP- incubated soils.

In soils incubated with ^{13}C-BZ, a smaller set of OTUs became highly abundant over the course of the incubation than was observed for ^{13}C-BP. By day one (early) dominant labeled OTUs in the ^{13}C-BZ incubation were *Methylobacterium jeotgali* (OTU 4), which comprised 19% of the total labeled sequences, and *Rhodococcus qingshengii* (OTU 8), which represented 11% of sequences. Other dominant labeled OTUs clustered with *Escherichia* spp. (OTU 10, 9%), *Afipia birgiae* (OTU 12, 3%), and *Salmonella enterica* (OTU 27, 3%). The late (four days) incubated

^{13}C-BZ soils were highly dominated by sequences clustering with *Burkholderia phytofirmans* (OTU 1, 34%), *Burkholderia sediminicola* (OTU 3, 27%), and *Cupriavidus basilensis* (OTU 6, 16%). *Burkholderia sordidicola* (OTU 18, 5%) and, similar to early ^{13}C-BP, *Arthrobacter oryzae* (OTU 2, 5%) were also dominantly labeled OTUs.

DISCUSSION

In this study, we sought to better understand microbial biodegradation of PCBs and how C from aromatic compounds flows through the soil microbial ecosystem. Using parallel SIP incubation studies, this experiment reveals how carbon flows from BP and BZ into a PCB-contaminated soil microbial community, and which different microbial taxa participate in these processes. The results revealed that a wide array of bacteria derived C from biodegradation of ^{13}C-BP and/or ^{13}C-BZ, either directly or indirectly through cross-feeding. The majority of taxa appeared to perform either BP or BZ utilization, but not both. Through this time-course study, the data also suggest a difference in ecological growth strategy for the two microbial guilds. A subset of organisms derived C from both substrates, which could be explained by three scenarios: a) they possess both BP and BZ pathways, and thus can mineralize both substrates, b) they are BZ utilizers that scavenge BZ released by BP-utilizers with only the upper pathway as they release BZ or c) they scavenge other substrates or dead biomass generated by BP utilizing microbes.

SIP showed that biotransformation of BP and BZ in this soil was mainly mediated by *Actinobacteria* and *Proteobacteria* (Table 2-3). Many of the genera detected appeared to derive carbon from either labeled BP, including unclassified *Acidimicrobiales, Pseudomonas, Marmoricola*, and *Arthrobacter*, or labeled BZ, including *Rhodococcus, Escherichia, Mesorhizobium, Propionibacterium, Burkholderia*, and *Cupriavidus*. However, we also detected

genera which derived carbon from both labeled BP and BZ including *Arthrobacter*, *Methylobacterium*, *Afipia*, *Gemmatimonas,* and *Rhodanobacter*. The presence of microbial taxa in both of the labeled fractions indicates that it is likely that some portion of the overlapping microbial population detected is capable of transforming BP and BZ, or completing the upper and lower pathways of BP degradation.

Pseudomonas and *Arthrobacter*, implicated in BP metabolism in this study, have often been reported to be associated with the upper pathway of BP degradation and are known to contain the *bph* operon and have been shown to degrade PCBs (9, 19, 32). Pseudomonads have previously been reported to be the most prevalent group of bacteria in the biodegradation of complex organic compounds (11). The unclassified *Acidimicrobiales* and *Marmoricola*, are previously unidentified in the metabolism of BP or PCBs. The closest RDP match for the unclassified *Acidimicrobiales* was *Terriglobus saanensis* (Table 2-3), a bacterium that has been previously isolated from tundra soils (33) and has wide potential physiological capabilities in soils, although the ecological role of many acidobacteria are poorly understood (34). *Nocardioides jensenii* was the closest RDP match for the unclassified *Marmoricola* (Table 2-3). In multiple studies, the genus *Nocardioides* has been previously identified as involved in degradation of PCB and other contaminants (18, 19, 35, 36).

Among organisms that derived carbon from BZ, the genera *Rhodococcus*, *Burkholderia*, and *Cupriavidus* have all previously been implicated in PCB degradation. *Propionibacterium* has been implicated in degradation of benzoate (37). Sequences clustering with *Escherichia* were also found in heavy labeled early (d01) BZ incubated soils and have been previously implicated in BZ degradation (11).

We further investigated the dominant, labeled OTUs in each of the ^{13}C-incubated incubated samples to characterize the major changes at each time point and with each substrate (Figure 2-2). The diversity and overlap of organisms that derived carbon from each substrate over time (Figure 2-1) shows how the guilds may interact (Figure 2-2). The early and late (d 04 and 14) ^{13}C-BP shared more OTUs between them than with ^{13}C-BZ incubations, with fewer unique OTUs present over time (Figure 2-1). In ^{13}C-BZ labeled populations, there were increasing numbers of unique OTUs over time suggesting that C flowed into the community via secondary feeding.

Labeled sequences from late ^{13}C-BP (d14) and early ^{13}C-BZ (d01) incubated soils were compared in an effort to determine if there were microorganisms present which could transition between the upper and lower pathways of BP degradation (Figure 2-1). When all of the OTUs present in ^{13}C-BP and ^{13}C-BZ incubated soils were compared, 16% of OTUs are shared between the two substrates, which indicates that the majority of the microbial populations do not utilize both the upper and lower BP degradation pathway; rather the pathways occur primarily in separate populations. Thirteen percent of the OTUs were shared between late ^{13}C-BP (d14) and early ^{13}C-BZ (d01), with more unique OTUs present in late ^{13}C-BP (d14) soils (14.8%). This increased number of OTUs in the late ^{13}C-BP (d14) soils could be due to diffusion of the late ^{13}C-signal through the community over time. The 12% shared OTUs between late ^{13}C-BP (d14) and early ^{13}C-BZ (d01) suggests there is a portion of the microbial populations which contain both the upper and lower pathways of PCB degradation (Figure 2-1). It could also indicate that some microorganisms from the later incubations of ^{13}C-BP incubated soils may have picked up the ^{13}C signal from degradation products of the upper portion of the pathway (22). The increase in the number of ^{13}C-associated OTUs observed over the course of the incubation period in the

BP-incubated soils suggests that carbon flowed through BP utilizers into other organisms, including BZ utilizers, in the soil community (Figure 2-1, Figure 2-2, Table 2-2). Microorganisms labeled in the late BP soils could also represent both fast and slow growing microorganisms, which is potentially why there are more unique OTUs in the later BP than early BP samples. The OTU overlap between late BP and early BZ time points also indicates that approximatly 12.6% of the microbes in the late BP time point are in fact utilizing BZ. At most, 16% of BP-degrading taxa may possess both pathways, but the number is likely lower given that cross-feeding on metabolites or dead biomass of BP-degraders may account for some portion of the overlap. The overlapping OTUs were mostly associated with the phylum Proteobacteria, which has been implicated in both BP and BZ degradation, but also contains known heterotrophs and potential cross feeders (11, 19).

The substrate utilization data ($^{13}CO_2$ evolution) provides one insight to the ecological growth strategies of the microorganisms associated with each of the substrates. The microcosms incubated with BZ had a much faster rate of $^{13}CO_2$ evolution than the BP-incubated microcosms (Table 2-1). This difference could be explained by the structural differences between BP and BZ, with BP having a more complex structure. It has been found that compounds that require a large number of enzymatic steps to be degraded, such as BP, may result in lowered substrate use efficiency (38).

Shifts in bacterial diversity were different for the two incubated substrates (Table 2-2). The effective number of species in ^{13}C-BZ incubated microcosms decreased over time (d01 to d04), however the other measure of diversity, Chao 1 index, increased over time. Although Chao 1 increased over time, when taking evenness measures into account (effective number of species, Simpson Index), the decrease could indicate that although there may have been fewer

species detected in the ^{13}C-BZ labeled populations over time, the species that were detected increased in abundance. In soils incubated with ^{13}C-BP, both the effective number of species and number of OTUs decreased over time (d04 to d14). Although there were fewer OTUs in BZ-incubated soils, overall the microcosms incubated with ^{13}C-BZ had a higher and faster rate of substrate utilization than the high number of BP utilizing OTUs. When coupled with a lower diversity in the BZ incubated samples and a higher amount of substrate used, this could indicate that the microorganisms associated with BP degradation are slower-growing populations.

The microbial communities associated with each of the substrates had very different population shifts. Recent work recognizing broad, ecologically-based, classes of bacteria (39–41) could also explain the differences in substrate use and ^{13}C-incoporation. These ecologically defined classes are similar to the r-k scheme (42) and defined as: 'copiotrophic' bacteria which are described as organisms which show an increase in numbers in response to the addition of labile carbon sources to soil or 'oligotrophic' bacteria which do not show a marked increase in numbers in response to the addition of a labile C source to the soil. It has been found through experimentation, observation, and meta-analytical approaches that Gram negative bacteria, such as members of the phylum Bacteroidetes and subphyla β- & γ-Proteobacteria, which rapidly colonize and grow under labile C-rich conditions, are copiotrophs. Gram positive soil bacteria, such as some Actinobacteria, tend to be more successful in resource-limited situations and are considered oligotrophs (39, 42). In addition, Fierer et al. (39) found that Acidobacteria can also be considered as oligotrophs.

The fast response to labeled BZ by the microbial community indicates that microbes involved in the lower pathway of BP degradation may be adapted to quickly react and utilize ^{13}C-BZ when it is present in the environment, indicative of a copiotrophic population. Whereas,

the initially lower rates of $^{13}CO_2$ evolution in ^{13}C-BP-incubated microcosms might indicate that the populations involved may be classified as "oligotrophic" and adapted for the optimal utilization of the recalcitrant C, such as BP, available (11). Of the five most abundant OTUs from both BP early and late incubations, half have been previously described as oligotrophic populations (39). Also, approximately 49.3% of the BP incubated population was gram positive, or associated with oligotrophic microorganisms, with 42.7% gram negative and 8% unknown. However in the BZ incubated soils, all but one of the top five most abundant OTUs from each time point have been described as copiotrophic microbes. In BZ incubated soils, 72.8% of the total microbial population was gram negative, or copiotrophic associated, with 23.8% gram positive and 3.4% unknown.

Overall, our data suggest that biotransformation of BP and BZ was mainly performed by microorganisms with different growth strategies. Stable isotope probing of contaminated soils where we used ^{13}C-BP and ^{13}C-BZ labeled substrates in parallel incubations over time reveals how aromatic compounds, such as PCBs, are processed by the microbial community. Our results suggest that the microbial populations associated with BZ are fast growing, copiotrophic organisms, while microbes which transform BP are oliotrophic, slower growing, organisms. The different patterns of use of the available carbon source could be an adaptation for the efficient use of a substrate or could indicate that the microbial community involved in the upper and lower pathway. Future, studies should focus on the stable isotope labeled metabolites produced by BP and BZ utilizing populations over time. By examining labeled substrates and metabolites over time, researchers can further understand how carbon flows between the different microbial guilds involved in PCB biodegradation.

ACKNOWLEDGMENTS

The project described was supported by Grant Number 5P20RR016466 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH) and a graduate fellowship from the Alaska Idea Network for Biomedical Research Excellence (INBRE), and EPSCoR NSF award #EPS-0701898 & the State of Alaska. MCL acknowledges the David L. Boren National Security Education Program (NSEP) Fellowship for support. OU acknowledges the support of Czech Science Foundation (grant no. 13-20414P).

We thank Jakub Rídl and Miluse Hroudová of the Institute of Molecular Genetics, Czech Academy of Sciences for help with sequencing. Thanks to Jacobs Engineering for access to soil samples. Finally, thanks to Roger Ruess, D. Lee Taylor, and Todd O'Hara for valuable feedback and comments to the manuscript.

REFERENCES

1. **Singer AC, Gilbert ES, Luepromchai E, Crowley DE**. 2000. Bioremediation of polychlorinated biphenyl-contaminated soil using carvone and surfactant-grown bacteria. Appl. Microbiol. Biotechnol. **54**:838–43.

2. **Waisner SA, Morrow AB, Nestler CC, Coyle C, Medina VF**. 2011. Chemical Degradation of PCBs in Alaskan Soils.

3. **Francova K, Macková M, Macek T, Sylvestre M**. 2004. Ability of bacterial biphenyl dioxygenases from Burkholderia sp. LB400 and Comamonas testosteroni B-356 to catalyse oxygenation of ortho-hydroxychlorobiphenyls formed from PCBs by plants. Environ. Pollut. **127**:41–48.

4. **Uhlik O, Jecna K, Mackova M, Vlcek C, Hroudova M, Demnerova K, Paces V, Macek T**. 2009. Biphenyl-metabolizing bacteria in the rhizosphere of horseradish and bulk soil contaminated by polychlorinated biphenyls as revealed by stable isotope probing. Appl. Environ. Microbiol. **75**:6471–7.

5. **Kuiper I, Lagendijk EL, Bloemberg G V, Lugtenberg BJJ**. 2004. Rhizoremediation: a beneficial plant-microbe interaction. Mol. Plant. Microbe. Interact. **17**:6–15.

6. **Leigh MB, Prouzová P, Macková M, Macek T, Nagle DP, Fletcher JS**. 2006. Polychlorinated biphenyl (PCB)-degrading bacteria associated with trees in a PCB-contaminated site. Appl. Environ. Microbiol. **72**:2331–42.

7. **Leewis M-C, Reynolds CM, Leigh MB**. 2013. Long-term effects of nutrient addition and phytoremediation on diesel and crude oil contaminated soils in subarctic Alaska. Cold Reg. Sci. Technol. **96**:129–137.

8. **Koubek J, Mackova M, Macek T, Uhlik O**. 2013. Diversity of chlorobiphenyl-metabolizing bacteria and their biphenyl dioxygenases in contaminated sediment. Chemosphere **93**:1548–55.

9. **Pieper DH, Seeger M**. 2008. Bacterial metabolism of polychlorinated biphenyls. J. Mol. Microbiol. Biotechnol. **15**:121–38.

10. **Pieper DH**. 2005. Aerobic degradation of polychlorinated biphenyls. Appl. Microbiol. Biotechnol. **67**:170–91.

11. **Uhlik O, Wald J, Strejcek M, Musilova L, Ridl J, Hroudova M, Vlcek C, Cardenas E, Mackova M, Macek T**. 2012. Identification of bacteria utilizing biphenyl, benzoate, and naphthalene in long-term contaminated soil. PLoS One **7**:e40653.

12. **Harwood CS, Parales RE**. 1996. The beta-ketoadipate pathway and the biology of self-identity. Annu. Rev. Microbiol. **50**:553–90.

13. **Gilbert ES, Crowley DE**. 1997. Plant compounds that induce polychlorinated biphenyl biodegradation by Arthrobacter sp. strain B1B. Appl. Environ. Microbiol. **63**:1933–8.

14. **Tillmann S, Strömpl C, Timmis KN, Abraham W-R**. 2005. Stable isotope probing reveals the dominant role of Burkholderia species in aerobic degradation of PCBs. FEMS Microbiol. Ecol. **52**:207–17.

15. **Ohtsubo Y, Nagata Y, Kimbara K, Takagi M, Ohta A**. 2000. Expression of the bph genes involved in biphenyl/PCB degradation in Pseudomonas sp. KKS102 induced by the biphenyl degradation intermediate, 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid. Gene **256**:223–8.

16. **Chen Y, Murrell JC**. 2010. When metagenomics meets stable-isotope probing: progress and perspectives. Trends Microbiol. **18**:157–63.

17. **Neufeld JD, Vohra J, Dumont MG, Lueders T, Manefield M, Friedrich MW, Murrell JC**. 2007. DNA stable-isotope probing. Nat. Protoc. **2**:860–6.

18. **Uhlik O, Leewis M-C, Strejcek M, Musilova L, Mackova M, Leigh MB, Macek T**. 2013. Stable isotope probing in the metagenomics era: A bridge towards improved bioremediation. Biotechnol. Adv. **31**:154–165.

19. **Leigh MB, Pellizari VH, Uhlík O, Sutka R, Rodrigues J, Ostrom NE, Zhou J, Tiedje JM**. 2007. Biphenyl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). ISME J. **1**:134–48.

20. **He R, Wooller MJ, Pohlman JW, Catranis C, Quensen J, Tiedje JM, Leigh MB**. 2012. Identification of functionally active aerobic methanotrophs in sediments from an arctic lake using stable isotope probing. Environ. Microbiol.

21. **Bell TH, Yergeau E, Martineau C, Juck D, Whyte LG, Greer CW**. 2011. Identification of Nitrogen-Incorporating Bacteria in Petroleum-Contaminated Arctic Soils Using 15N DNA-SIP and Pyrosequencing. Appl. Environ. Microbiol. **77**:4163–4171.

22. **DeRito CM, Pumphrey GM, Madsen EL**. 2005. Use of field-based stable isotope probing to identify adapted populations and track carbon flow through a phenol-degrading soil microbial community. Appl. Environ. Microbiol. **71**:7858–65.

23. **Dumont MMG, Murrell JC**. 2005. Stable isotope probing - linking microbial identity to function. Nat. Rev. Microbiol. **3**:499–504.

24. **Engineers AC of**. 2003. Final Clean-Up Decision Document: Drury Gulch Kodiak Island, Alaska; July 2003Order A Journal On The Theory Of Ordered Sets And Its Applications.

25. **Slater H, Gouin T, Leigh MB**. 2011. Assessing the potential for rhizoremediation of PCB contaminated soils in northern regions using native tree species. Chemosphere **84**:199–206.

26. **He R, Wooller MJ, Pohlman JW, Quensen J, Tiedje JM, Leigh MB**. 2012. Shifts in identity and activity of methanotrophs in arctic lake sediments in response to temperature changes. Appl. Environ. Microbiol. **78**:4715–23.

27. **He R, Wooller MJ, Pohlman JW, Quensen J, Tiedje JM, Leigh MB**. 2012. Diversity of active aerobic methanotrophs along depth profiles of arctic and subarctic lake water column and sediments. ISME J. **6**:1937–48.

28. **Engelbrektson A, Kunin V, Wrighton KC, Zvenigorodsky N, Chen F, Ochman H, Hugenholtz P**. 2010. Experimental factors affecting PCR-based estimates of microbial species richness and evenness. ISME J. **4**:1–6.

29. **Schloss PD, Gevers D, Westcott SL**. 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One **6**:e27310.

30. **Bailey RE, Gonsior SJ, Rhinehart WL**. 1983. Biodegradation of the monochlorobiphenyls and biphenyl in river water. Environ. Sci. Technol. **17**:617–621.

31. **Jost L**. 2006. Entropy and diversity. Oikos **2**.

32. **Erickson BD, Mondello FJ**. 1992. Nucleotide sequencing and transcriptional mapping of the genes encoding biphenyl dioxygenase, a multicomponent polychlorinated-biphenyl-degrading enzyme in Pseudomonas strain LB400. J. Bacteriol. **174**:2903–12.

33. **Männistö MK, Rawat S, Starovoytov V, Häggblom MM**. 2011. Terriglobus saanensis sp. nov., an acidobacterium isolated from tundra soil. Int. J. Syst. Evol. Microbiol. **61**:1823–8.

34. **Eichorst S a, Breznak J a, Schmidt TM**. 2007. Isolation and characterization of soil bacteria that define Terriglobus gen. nov., in the phylum Acidobacteria. Appl. Environ. Microbiol. **73**:2708–17.

35. **Fetzner S**. 2012. Ring-cleaving dioxygenases with a cupin fold. Appl. Environ. Microbiol. **78**:2505–14.

36. **Liang Y, Van Nostrand JD, Deng Y, He Z, Wu L, Zhang X, Li G, Zhou J**. 2011. Functional gene diversity of soil microbial communities from five oil-contaminated fields in China. ISME J. **5**:403–13.

37. **Thorenoor N, Kim Y-H, Lee C, Yu M-H, Engesser K-H**. 2009. A previously uncultured, paper mill Propionibacterium is able to degrade O-aryl alkyl ethers and various aromatic hydrocarbons. Chemosphere **75**:1287–93.

38. **Manzoni S, Taylor P, Richter A, Porporato A, Agren GI**. 2012. Environmental and stoichiometric controls on microbial carbon-use efficiency in soils. New Phytol. **196**:79–91.

39. **Fierer N, Bradford M a, Jackson RB**. 2007. Toward an ecological classification of soil bacteria. Ecology **88**:1354–64.

40. **Eilers KG, Lauber CL, Knight R, Fierer N**. 2010. Shifts in bacterial community structure associated with inputs of low molecular weight carbon compounds to soil. Soil Biol. Biochem. **42**:896–903.

41. **Pumphrey GM, Madsen EL**. 2008. Field-based stable isotope probing reveals the identities of benzoic acid-metabolizing microorganisms and their in situ growth in agricultural soil. Appl. Environ. Microbiol. **74**:4111–8.

42. **Margesin R, Hämmerle M, Tscherko D**. 2007. Microbial activity and community composition during bioremediation of diesel-oil-contaminated soil: effects of hydrocarbon concentration, fertilizers, and incubation time. Microb. Ecol. **53**:259–69.

FIGURES

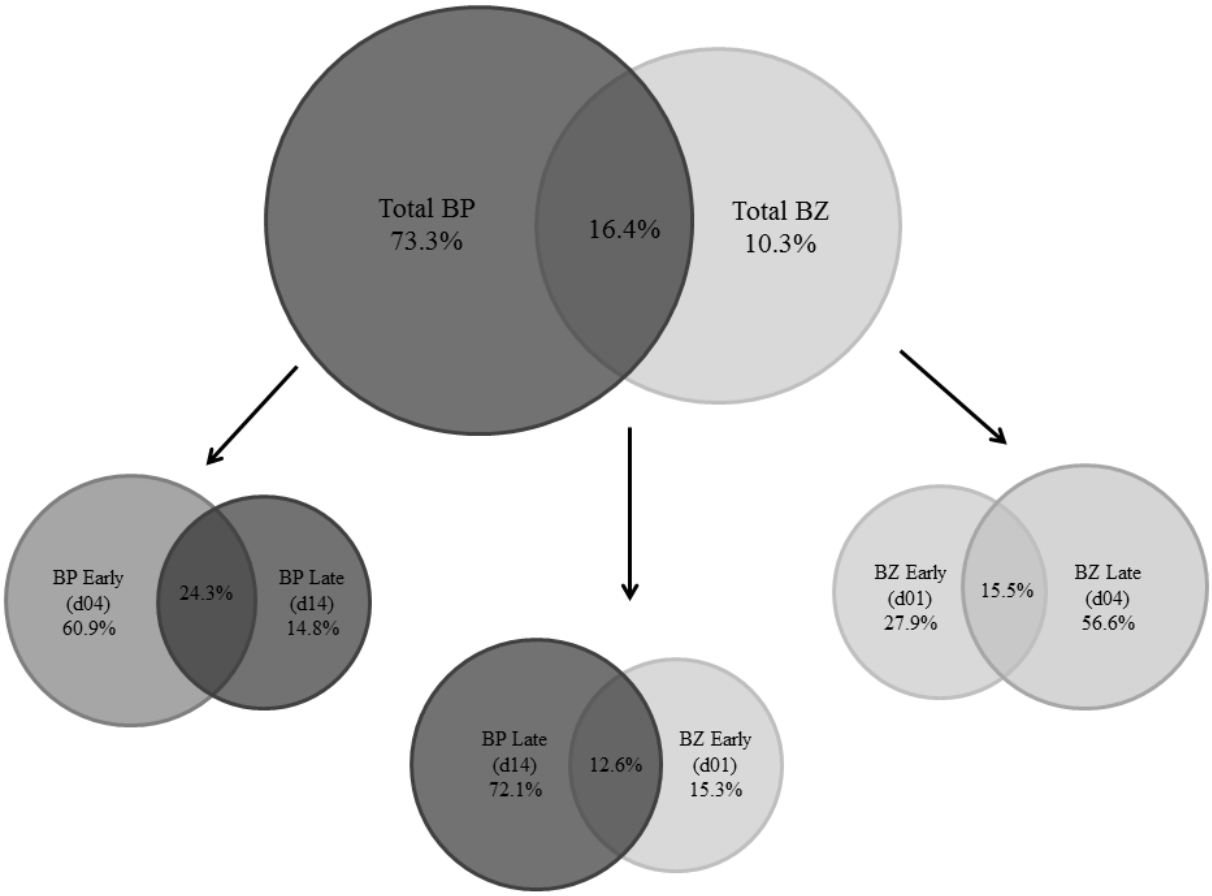

Figure 2-1. Venn diagraphs showing the overall overlap of taxonomic units between the two labeled substrates at different time points and the total labeled communities from each substrate. The percentages represent the percent of OTUs unique to each of the individual time points and/or substrate. Sequences were grouped at 3% dissimilarity.

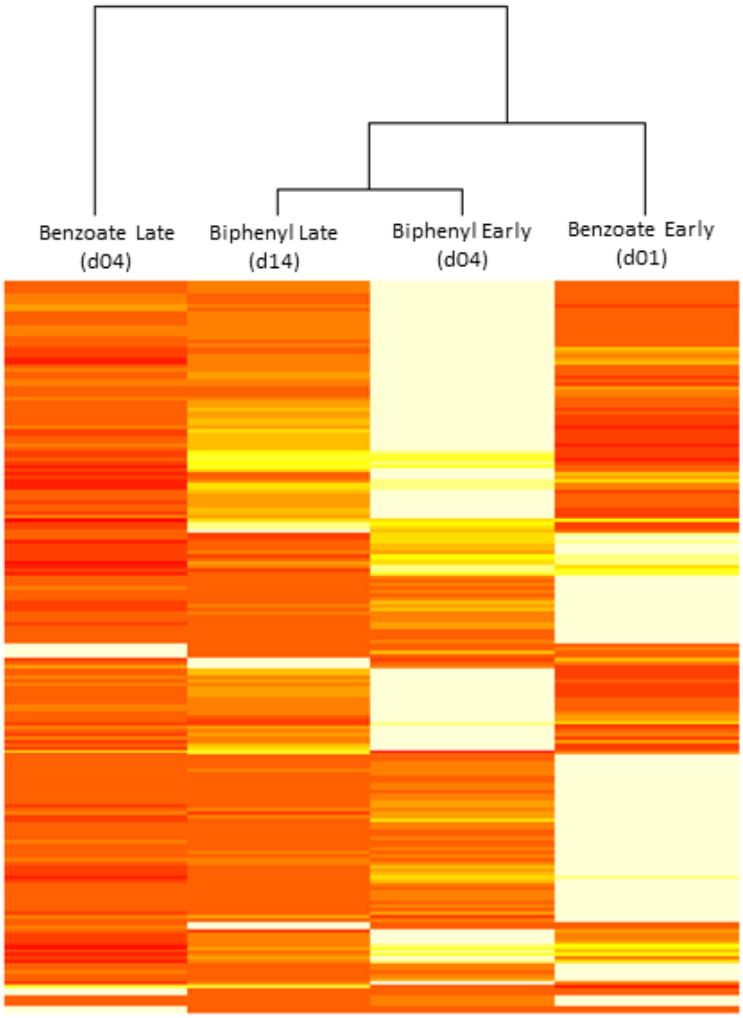

Figure 2-2. Heat map generated using relative abundance and distribution of the top 200 OTUs of heavy labeled 16S rRNA genes from each sample. The colour code indicates relative abundance, ranging from light yellow (low abundance) to dark red (high abundance). White boxes indicated that no reads were assigned to the corresponding OTU. The dendrogram was generated using complete linkage clustering based on a Euclidian distance measure.

TABLES

Table 2-1. Mineralization of ^{13}C-BP and ^{13}C -BZ as determined by $^{13}CO_2$ evolution. ^{13}C-BP and ^{13}C –BZ mineralization calculated as a range assuming 60-80% of BP utilized is mineralized to CO_2 (30).

SIP incubation time (days)	$^{13}CO_2$ excess in headspace (nmol)		^{13}C substrate mineralized (ug)	
	Benzoate	Biphenyl	Benzoate	Biphenyl
1	31.10	13.99	0.54 - 0.72	0.24 - 0.32
4	416.57	20.31	7.2 - 9.6	0.35 - 0.47
14	725.70	45.80	12.5 - 16.7	0.79 - 1.05

Table 2-2. Diversity estimates and number of sequences obtained after raw sequence data processing.

Sample ID		Sequences after processing	# OTUs (0.03)	Chao1	1/D (Inverse Simpson)	Effective Number of Species
Biphenyl	Day 04	12529	2041	3198	172.41	678.74
	Day 14	5684	938	1348	44.47	268.64
Benzoate	Day 01	13474	310	553	16.57	54.05
	Day 04	14848	515	856	4.57	10.87
Total Community		9745	2174	4876	159.51	717.34

Table 2-3. Top OTUs detected in ^{13}C-DNA after incubation of soil with ^{13}C-biphenyl and ^{13}C-benzoate.

Phylogenetic affiliations	*Closest RDP SeqMatch Type Strain (Similarity score)*	*Accession No.*	OTU	Biphenyl		Benzoate		Total Community
				Day 04	Day 14	Day 01	Day 04	
Acidobacteria								
unclassified	Terriglobus saanensis SP1PR4 (0.901)	HM214537	22	1 (121)	1 (32)	- (64)	- (26)	1 (135)
Actinobacteria								
Actinobacteria	Aciditerrimonas ferrireducens IC-180 (0.896)	AB517669	31	<0.5 (36)	<0.5 (25)	<0.5 (14)	<0.5 (6)	1 (119)
Actinobacteria								
Actinobacteria								
Micrococcaceae								
	Arthrobacter oryzae KV-651 (1.000)	AB279889	2	1 (187)	12 (683)	1 (93)	5 (682)	2 (201)
	Arthrobacter oryzae KV-651 (0.995)	AB279889	26	<0.5 (26)	2 (90)	- -	<0.5 (32)	<0.5 (30)
	Arthrobacter humicola KV-653 (1.000)	AB279890	94	<0.5 (18)	- -	- -	<0.5 (48)	<0.5 (16)
Nocardiaceae								
	Rhodococcus qingshengii djl-6 / R. jialingiae djl-6-2 (1.000)	DQ090961 / DQ185597	8	<0.5 (46)	<0.5 (6)	11 (1418)	<0.5 (9)	- -
	Nocardioides jensenii KCTC 9134 (0.975)	AF005006	13	2 (197)	1 (62)	1 (81)	<0.5 (11)	2 (150)
Propionibacteriaceae								
	Propionibacterium acnes ATCC6919 (1.000)	AB042288	25	<0.5 (14)	<0.5 (19)	2 (320)	<0.5 (30)	<0.5 (1)
Actinobacteria								
unclassified	Nocardioides luteus KCTC 9575 (0.797)	AF005007	11	2 (189)	1 (42)	<0.5 (58)	<0.5 (17)	1 (145)
Actinobacteria								
EB1017	Aciditerrimonas ferrireducens IC-180 (0.896)	AB517669	19	2 (197)	<0.5 (16)	<0.5 (66)	<0.5 (31)	1 (114)
AD3								
ABS-6								
unclassified	Acidothermus cellulolyticus 11B (0.842)	CP000481	5	4 (550)	3 (175)	1 (180)	- (30)	4 (419)
Gemmatimonadetes								
Gemmatimonadetes								
Gemmatimonadaceae								
	Gemmatimonas aurantiaca T-27 (0.891)	AB072735	9	2 (287)	2 (110)	1 (100)	<0.5 (43)	3 (283)
	Gemmatimonas aurantiaca T-27 (0.896)	AB072735	15	1 (158)	1 (47)	<0.5 (55)	<0.5 (29)	2 (171)
Proteobacteria								
Alphaproteobacteria								
Methylobacteriaceae								
	Methylobacterium jeotgali S2R03-9 (1.000)	DQ471331	4	1 (135)	<0.5 (10)	19 (2552)	- -	- -
Bradyrhizobiaceae								
	Afipia birgiae 34632 / A. broomeae F186 (1.000)	AF288304 / U87759	12	1 (149)	1 (34)	3 (417)	<0.5 (32)	2 (166)
Phyllobacteriaceae								
	Mesorhizobium australicum WSM2073 (0.990)	AY601516	32	<0.5 (8)	- -	3 (365)	- -	- -
Methylobacteriaceae								
	Methylobacterium fujisawaense DSM 5686 / M. oryzae CBMB20 (0.995)	AJ250801 / AY683045	38	<0.5 (14)	<0.5 (2)	2 (332)	- -	- -
unclassified								
	Afipia birgiae 34632 (0.960)	AF288304	56	<0.5 (10)	- -	1 (176)	- -	- -

Table 2-3 continued

Betaproteobacteria											
Burkholderiaceae											
	Burkholderia sediminicola HU2-65W (1.000)	EU035613	3	-	-	<0.5	(21)	-	-	27	(4036)	-	-
	Burkholderia sordidicola SNU 020123 (1.000)	AF512827	18	<0.5	(2)	<0.5	(2)	-	-	5	(795)	-	-
	Burkholderia ginsengisoli KMY03 (0.980)	AB201286	193	-	-	-	-	-	-	1	(78)	-	-
	Burkholderia phytofirmans PsJN (1.000)	AY497470	1	-	-	<0.5	(20)	-	-	34	(5039)	-	-
	Burkholderia sordidicola SNU 020123 (0.985)	AF512827	252	-	-	-	-	-	-	<0.5	(63)	-	-
	Burkholderia soli GP25-8 (0.990)	DQ465451	259	-	-	-	-	<0.5	(1)	<0.5	(53)	-	-
	Cupriavidus basilensis DSM 11853 (1.000)	AF312022	6	-	-	<0.5	(6)	-	-	16	(2335)	-	-
unclassified	Schlegelella thermodepolymerans K14 (0.920)	AY152824	14	2	(221)	2	(90)	1	(141)	<0.5	(26)	1	(110)
Gammaproteobacteria											
Sinobacteraceae	Marinobacter excellens KMM 3809 (0.891)	AY180101	21	1	(145)	1	(52)	1	(116)	-	(12)	1	(88)
Enterobacteriaceae											
	Salmonella enterica LT2	AE006468	27	<0.5	(14)	<0.5	(4)	3	(384)	<0.5	(11)	-	-
	Escherichia/Shigella dysenteriae / Escherichia/Shigella flexneri	X96966 / X96963	10	<0.5	(38)	1	(33)	9	(1154)	<0.5	(21)	-	-
Pseudomonadaceae											
Pseudomonadaceae											
	Pseudomonas lutea OK2 (0.995)	AY364537	24	-	-	4	(246)	<0.5	(3)	<0.5	(9)	-	-
	Pseudomonas asplenii ATCC 23835T (0.995)	AB021397	173	-	-	1	(73)	<0.5	(7)	-	-	-	-
Xanthomonadaceae											
	Rhodanobacter soli DCY45 (1.000)	FJ605268	17	<0.5	(16)	2	(115)	<0.5	(56)	1	(101)	-	-
	Stenotrophomonas maltophilia IAM 12423 (1.000)	AB294553	120	<0.5	(3)	1	(68)	<0.5	(30)	<0.5	(3)	-	-
TM7											
TM7-3											
unclassified	Thermosipho geolei DSM 13256 (0.796)	AJ272022	7	-	-	5	(264)	-	-	<0.5	(2)	<0.5	(1)

Taxonomic assignments were generated using the Ribosomal Database Project's classifier.

aThe numbers of pyrosequencing reads assigned using the RDP classifier (80% confidence threshold) are shown in parentheses.

b sequences represented < 0.5% of total sequences present in the group

c Results are to the closest match; similarity score shown in parentheses

SUPPLEMENTAL FIGURES

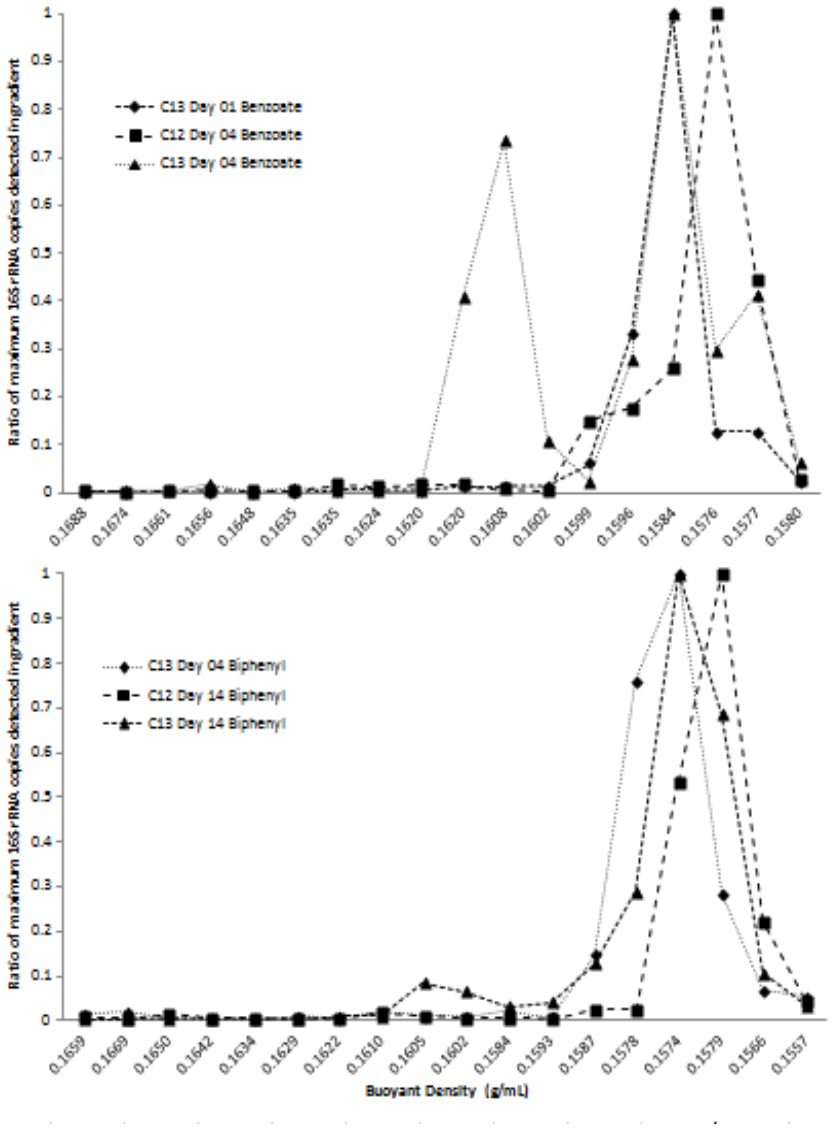

Figure S2-1. Detection of ^{13}C-DNA in density gradient fractions derived from DNA in (a) ^{13}C – BZ incubated soils at 1 day, 4 days, and ^{12}C -BZ incubated control soils at day 04; (b) and ^{13}C-BP incubated soils at 4 days, 14 days, and ^{12}C -BP incubated control soils at day 14.

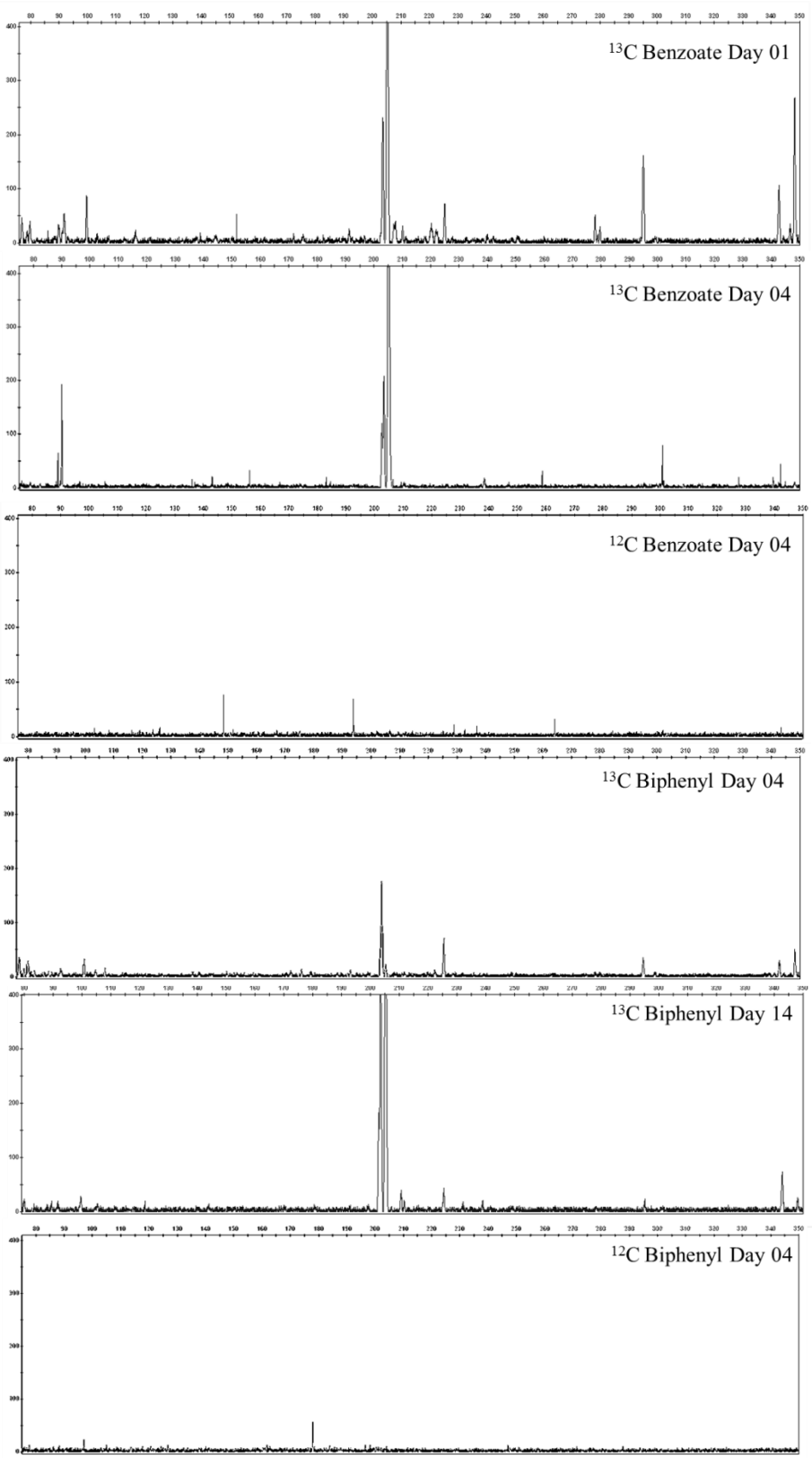

Figure S2-2. T-RFLP profiles of 16S rRNA gene amplicons digested with *HhaI*.

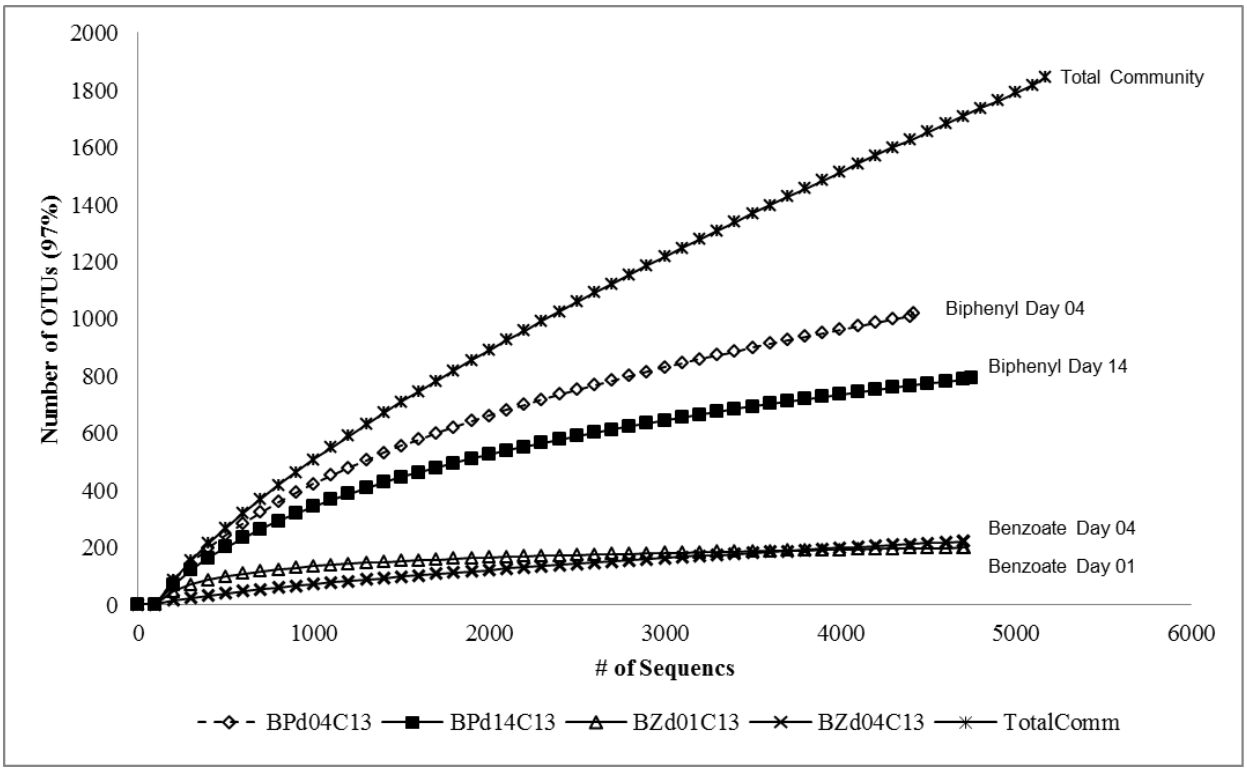

Figure S2-3. Rarefaction curves for the numbers of pyrosequencing reads and operational taxonomic units (OTUs) at 3% dissimilarity.

SUPPLEMENTAL TABLE

Table S2-1. Relative abundance of total reads and phylogenetic affiliations of all pyrosequencing reads from ^{13}C-DNA obtained from SIP of PCB contaminated soil. The numbers of pyrosequencing reads assigned using the RDP classifier (80% confidence threshold) are shown in parentheses. Sequences represented < 0.5% of total sequences present in the group were represented as "-".

Phylogenetic affiliations	Biphenyl				Benzoate				Total Community	
	Day 04		Day 14		Day 01		Day 04			
Acidobacteria										
Acidobacteria										
Acidobacteriaceae	1.53	(175)	1.28	(69)	1.03	(134)	-	(16)	2.69	(238)
unclassified	1.78	(203)	1.71	(92)	2.23	(290)	-	(28)	1.87	(166)
Chloracidobacteria										
unclassified	0.89	(102)	0.48	(26)	-	(27)	-	(19)	1.91	(169)
Solibacteres										
Solibacteraceae	4.14	(473)	3.28	(177)	0.91	(119)	-	(52)	6.04	(535)
unclassified										
unclassified	2.83	(323)	1.93	(104)	2.07	(270)	-	(65)	4.69	(415)
others	4.28	(489)	2.75	(148)	-	(40)	-	(50)	3.15	(279)
Actinobacteria										
Actinobacteria										
Corynebacteriaceae	-	(10)	-	(0)	0.47	(61)	-	(1)	-	(0)
Intrasporangiaceae	1.26	(144)	1.78	(96)	0.94	(122)	-	(30)	0.98	(87)
Micrococcaceae	2.65	(303)	15.28	(824)	1.26	(164)	5.48	(808)	3.56	(315)
Nocardiaceae	0.62	(71)	-	(7)	11.15	(1451)	-	(10)	-	(27)
Nocardioidaceae	1.82	(208)	1.24	(67)	0.62	(81)	-	(12)	1.72	(152)
Propionibacteriaceae	0.12	(14)	-	(19)	2.46	(320)	-	(30)	-	(1)
Pseudonocardiaceae	0.26	(30)	0.20	(11)	0.94	(122)	-	(3)	-	(32)
unclassified	7.62	(871)	4.77	(257)	3.48	(472)	-	(65)	7.51	(665)
others	7.75	(885)	5.06	(273)	3.96	(515)	0.52	(77)	6.50	(576)
Bacteroidetes										
Sphingobacteria										
unclassified	0.21	(24)	-	(19)	-	(0)	-	(0)	0.50	(44)
Chloroflexi										
Ktedonobacteria										
unclassified	0.19	(22)	-	(4)	-	(0)	-	(1)	0.67	(59)
others	5.83	(666)	4.16	(224)	2.10	(274)	-	(59)	3.65	(323)
Firmicutes										
Clostridia										
Peptococcaceae	-	(3)	-	(0)	0.51	(66)	-	(1)	-	(1)
Bacilli										
Staphylococcaceae	-	(2)	-	(2)	0.58	(76)	-	(2)	-	(0)
Gemmatimonadetes										
Gemmatimonadetes										
Gemmatimonadaceae	5.99	(684)	4.93	(266)	1.52	(198)	-	(104)	7.46	(661)
unclassified	1.72	(197)	0.91	(49)	0.25	(32)	-	(7)	0.70	(62)

Table S2-1 continued

Nitrospirae										
	Nitrospira									
		Nitrospiraceae	0.85	(97)	0.52	(28)	0.23	(30)	-	(9)	1.58	(140)
Proteobacteria												
	Alphaproteobacteria											
		Bradyrhizobiaceae	1.36	(155)	0.69	(37)	3.54	(461)	-	(36)	1.95	(173)
		Hyphomicrobiaceae	1.30	(148)	0.89	(48)	-	(55)	-	(24)	1.94	(172)
		Methylobacteriaceae	1.30	(149)	-	(12)	22.17	(2886)	-	(0)	-	(1)
		Phyllobacteriaceae	0.18	(21)	-	(7)	2.82	(367)	-	(0)	-	(18)
		Sphingomonadaceae	0.60	(69)	-	(13)	1.87	(244)	-	(7)	-	(17)
		unclassified	1.295	(148)	0.538	(48)	1.813	(288)	-	(18)	1.5583	(138)
	Betaproteobacteria											
		Burkholderiaceae	-	(2)	0.96	(52)	-	(2)	84.34	(12430)	-	(5)
		Comamonadaceae	0.39	(45)	0.80	(43)	-	(42)	-	(15)	-	(32)
		Oxalobacteraceae	0.23	(26)	-	(11)	-	(0)	-	(50)	0.75	(66)
		unclassified	8.088	(924)	6.288	(362)	3.003	(400)	0.74	(114)	7.8026	(691)
	Deltaproteobacteria											
		Geobacteraceae	0.81	(92)	0.82	(44)	0.99	(129)	-	(8)	0.94	(83)
		Haliangiaceae	0.99	(113)	-	(19)	-	(18)	-	(3)	-	(41)
		Myxococcaceae	-	(36)	0.48	(26)	-	(0)	-	(1)	-	(20)
		Syntrophobacteraceae	0.78	(89)	0.46	(25)	-	(0)	-	(23)	1.00	(89)
	Gammaproteobacteria											
		Enterobacteriaceae	0.46	(52)	0.76	(41)	11.85	(1543)	-	(32)	-	(1)
		Pseudomonadaceae	-	(11)	6.66	(359)	1.11	(144)	-	(33)	-	(0)
		Sinobacteraceae	2.63	(301)	2.13	(115)	0.90	(117)	-	(37)	2.82	(250)
		Xanthomonadaceae	1.07	(122)	4.34	(234)	0.90	(117)	0.80	(118)	0.67	(59)
	unclassified											
		unclassified	1.15	(131)	0.93	(50)	-	(39)	-	(21)	0.98	(87)
TM7												
	TM7-3		0.00	(0)	5.31	(286)	-	(0)	-	(2)	-	(1)
Others			8.54	0.00	4.79	0.00	4.00	0.00	0.62	0.00	6.82	0.00
unclassified			8.54	(976)	4.79	(258)	4.00	(521)	0.62	(91)	6.82	(604)
Total			100	(11425)	100	(5391)	100	(13019)	100	(14738)	100	(8856)

Taxonomic assignments were generated using the Ribosomal Database Project's classifier.

Chapter 3

Microbial Community Structure and Functional Potential Associated With Four Boreal Forest Vegetation Types [1]

ABSTRACT

What drives microbial community structure and function is a fundamental question in microbial ecology. Many studies have found that plant successional stage and soil environmental factors affect the soil microbial community. As newer methods emerge (e.g. functional gene microarrays, next-generation sequencing), more comprehensive analyses of community structure and functional potential are possible. We used bar-coded 454-pyrosequencing of bacterial 16S rRNA genes and GeoChip functional gene microarray analyses to screen the microbial structural and functional potential of organic horizon soils associated with forest stands dominated by either *Picea glauca* (white spruce), *Picea mariana* (black spruce), *Populus tremuloides* (quaking aspen), or *Betula neoalaskana* (Alaskan paper birch), located within the Bonanza Creek Long Term Ecological Research Site (BNZ LTER) in interior Alaska. Results indicate that microbes with the genetic functional potential to degrade chlorinated solvents, hydrocarbons, herbicides, pesticides, and other aromatic compounds are present in varying amounts in the organic horizon of all stand types even though the sites are located in pristine forests. Additionally, the numbers and amounts of genes associated with carbon cycling differ among stand types. Bacterial community structure differed between forest stand types, and was influenced to a lesser extent by soil environmental factors. The majority of

[1] Leewis, M.C., D. L. Taylor, M.B. Leigh. Microbial community structure and functional potential associated with four boreal forest vegetation types. Prepared for submission to Applied and Environmental Microbiology.

16S rRNA gene sequences retrieved from the four different vegetation types clustered with Acidobacteria, Actinobacteria and Proteobacteria – together these three phyla represented 83% to 97% of all sequences. This study provides insight into the role of plants in driving microbial community structure and functional potential, and suggests that resilience to contaminants and phytoremediation potential exists in Alaskan boreal forest soils.

Subject Category: Microbial ecology and functional diversity of natural habitats

Keywords: Alaska / Boreal Forest / functional gene array / microbial community structure / soil / rhizosphere

INTRODUCTION

Microorganisms play a critical role in biogeochemical cycling and are drivers of soil ecosystem services and functioning, meditating in approximately 80-90% of soil biogeochemical processes (1). A fundamental question in microbial ecology concerns what factors drive microbial community structure and function. Microbial community structure refers to the diversity and composition of the community, while microbial function describes the metabolic capabilities active (or potentially active) in the community and ecosystem processes provided by the community (2, 3). Many environmental factors have been found to have an effect on microbial community function and structure, including; pH (4, 5), soil moisture (6), dominant litter type (7), and both xenobiotic and naturally occurring but anthropogenically released contaminants (8–11). One powerful determinant of microbial function in soils is vegetation type (12, 13). Because plants are diverse in their chemical composition, they impact the quantity and quality of carbon (C) and nutrients available to microbes through root turnover, root leachates, litter decay, as well as by altering pH and soil microclimate (moisture, temperature). These factors select for microbial communities which have functional similarities consistent within vegetation types (14, 15).

The northern boreal forest biome covers 11% of global land surface, contains 28% of the worlds' soil C, and houses 16% of the C stock sequestered in soils (16, 17). As global temperatures rise in response to climate warming, boreal forest regions will experience rapid changes due permafrost thawing and changing fire regimes (18). It is unknown how these changing climate dynamics may affect boreal soil microbial community functions, such as C sequestration and emissions, or the potential effects on the large soil C stores, due to direct effects on soil microbes or indirect effects via shifts in vegetation composition. Many studies

have found that with increasing temperature and nitrogen (N) deposition, there is increased loss of soil C through microbial activities and changes in the microbial community as measured by phospholipid fatty acid (PLFA) (19–21). Other studies of the boreal forest have found that tree stand age and dominant tree species also result in changes in fungal community structure (22–24). However there is a paucity of research which investigates both the microbial (bacterial) community structure and total community function between different boreal forest vegetation types or successional stages.

The Alaskan boreal region is an active location of petroleum development and transport, including aging and newly proposed oil pipelines, extensive use of diesel fuel oils for heat and increased rates of atmospheric deposition of PCBs and other pollutants at high latitudes, and legacies of contamination from formerly used defense sites. Contaminant remediation can be prohibitively expensive and logistically impractical in Alaska due to limited infrastructure such as roads. Phytoremediation is a potentially inexpensive and effective alternative method for soil contaminant cleanup (25, 26). The basis for the approach is that plant root systems can promote the growth of contaminant degrading bacteria through release of metabolic products, improved aeration and other mechanisms. Some of the aromatic compounds released by plant roots systems resemble pollutants or an intermediary in the metabolism of pollutants, and these intermediates have the potential to induce the biological degradation of petroleum compounds (27–29). Alaskan boreal vegetation is particularly attractive for investigations into microbial degradation of aromatic compounds due to the unusually high concentrations of secondary compounds present in many Alaskan plant species, which are thought to provide defenses against herbivory or protection from UV radiation (30, 31). Lab (32, 33), pot (34) and field (35) studies have demonstrated that native Alaskan vegetation has the potential to alter contaminant

biodegradation rates and/or microbial community structure in contaminated soils. Understanding the controls over microbial community structure and functional potential in boreal forest soils may advance phytoremediation technologies and is critical to understanding the potential effects of changes in climate, disturbance and vegetation on nutrient cycling and ecosystem function in the boreal forest biome.

The main objective of this study was to understand how vegetation affects microbial community and functional potential. We sought i) to determine if genes for organic pollutant degradation are present in uncontaminated soils ii) to understand the potential effects of dominant vegetation on microbial biogeochemical cycling, and iii) to quantify and compare microbial populations. To attain these objectives, pyrosequencing of 16S rRNA gene amplicons was used to characterize bacterial diversity and community composition, and functional genetic potential was assessed using a functional gene microarray (GeoChip 3.0; 35) This study used soils collected from the Bonanza Creek Long Term Ecological Research site (BNZ LTER), which have been previously investigated for fungal community structure and resilience (24, 37). We hypothesized that dominant vegetation type controls microbial community structure and function either through direct impacts of plant root exudates and detritus or indirectly through the influence of plants on soil chemistry, composition, and structure. We also hypothesize that certain plants foster increased contaminant degradation potential, even in the absence of substantial contamination.

MATERIALS AND METHODS

Study Site. All sites and soils used were also part of a previous study conducted on fungal community structure by Tayleo et al. (24). Soil samples were collected in 2008 from two upland early successional sites, three upland late successional sites, and three lowland late succession, black spruce dominated sites. These sites and plots are dealt with in more detail in Taylor et al. 2014 (37). The early successional sites are dominated by *Betula neoalaskana* (Paper Birch, UP1B) or *Populus tremuloides* (Aspen, UP1C), with some understory white spruce and are approximately 23-25 years old. The mature sites are dominated by large *Picea glauca* (White Spruce, UP3A, UP3B, UP3C) trees, have occasional birch and are approximately 225-230 years old. All upland sites are core monitoring sites within the Bonanza Creek LTER. The three lowland sites were established by Hollingsworth et al. (38) and are dominated by *Picea mariana* (Black Spruce, TKN0001, TKN0012, TKN0022). Vegetation data, collected by Hollingswort et al. (38), using the Braun-Blanquét cover-abundance scale, can be found in table S3-1, environmental chemistry can be found in table S3-2. Soils for molecular analyses were collected as follows: fifty soil cores were collected at 10 m intervals arrayed along four parallel transects from each of the sites. Cores were 1.8 cm diameter and collected to 20cm depth. Subsamples of 1 g were excised from each core and pooled for each site by horizon. Soils from the O horizon were used in this study. Soils were lyophilized then kept at -20°C until analysis. For further details on study sites and sample collection see Taylor et al. 2010 (24), 2014 (37)and Hollingsworth et al. 2006 (38), additional vegetation data are available on the Bonanza Creek website at www.lter.uaf.edu/.

Analysis of Nutrients in Soils. Soil physical and chemical analyses were conducted at the University of Alaska Fairbanks, Palmer Center for Sustainable Living. Properties assessed

were: pH, % loss on ignition (C), CEC, KCl- extractable $NO_3^-$and NH_4^+, and total P and K (39–43).

DNA Extractions and 16S rRNA gene pyrosequencing. Genomic DNA was extracted from 1-5 g of soil, for details see (37). Amplification and pyrosequencing were performed by a modification of a method described previously by Engelbrektson et al. (2010). A portion of the 16S rRNA gene (positions 8 to 357 [V1 and V2], *Escherichia coli* numbering) was amplified using a 8F primer with the addition of a Roche 454 A pyrosequencing adaptor (underlined) and a 10 base pair unique bar code (see supplemental materials) and 357R primer to which a Roche 454 B pyrosequencing adaptor (underlined) had also been added: 8F 5'-CGTATCGCCTCCCTCGCGCCATCAG-NNNNNNNNNN-AGAGTTTGATCMTGGCTCAG -3', 357R 5'- CTATGCGCCTTGCCAGCCCGCTCAGCTGCTGCSYCCCGTAG -3'. The MID bar code for each sample was unique to facilitate sorting of sequences from a single pyrosequencing run.

The PCR mixture (final volume, 25 µl) contained 1 µl each primer (10 µM), 0.5 µl dNTP mix (10 mM), 2.5 µl FastStart 10 X Buffer #2, 0.25 µl FastStart HiFi Polyumerase (5 U/µl), 18.75 µl molecular biology grade water (Roche). The following thermal cycling conditions were used: initial denaturation at 95°C for 3 min and 25 or 30 cycles of denaturation at 95°C for 30 sec, annealing for 1 min at 55°C, and extension at 72°C for 1.5 min, followed by a final extension period at 72°C for 10 min. Each PCR product was obtained in three parallel experiments, the resulting preparations were mixed, purified using Pure-Link PCR purification kit (Invitrogen, USA) and pooled for downstream sequencing. Roche 454 GS FLX Titanium sequencing (454 Life Sciences, USA) was performed on pooled reactions at the Institute of Molecular Genetics, Czech Academy of Sciences, Prague, Czech Republic.

Processing of Pyrosequencing Data. Raw pyrosequencing data were processed using the mothur software package version 1.30.2 (45) and following the standardized operating procedure detailed on the mothur website (accessed August 2013). In brief, sequences were discarded if there was more than one mismatch in the bar code, more than two mismatches in the primer sequence, more than eight homopolymers, or were less than 200 bp. Chimeras were detected and removed in mothur using the Uchime program (45). Sequences were normalized across samples by setting the number of sequences in the smallest group, and subsampling. The representative sequences have been deposited to the NCBI Short Read Archive under the accession number TBA.

Quantitative PCR (qPCR). To quantify bacterial and fungal DNA in soil samples, qPCR was performed on each sample according to He et al. (46, 47). Universal 16S rRNA gene primers 1108F (5'-ATGGYTGTCGTCAGCTCGTG-3') and 1132R (5'-GGGTTGCGCTCGGTTGC-3') (48) were used for bacterial qPCR while primers ITS1F (5'-TCCGTAGGTGAACCTGCGG-3') (49) and 5.8S (5'-TCGATGAAGAACGCAGCG-3') (50) were used for fungal qPCR. Relative quantity of bacterial or fungal DNA was determined in relation to standard curves (10 ng, 1 ng, 0.1 ng and 0.01 ng), consisting of genomic DNA extracted from pure cultures of *Pseudomonas stutzeri* JM300 for bacteria (51) or *Saccharomyces cerevesiae* for fungi.

GeoChip microarray hybridization. Three aliquots of 50 ng of DNA from each sample was amplified using the TempliPhi kit (GE Healthcare) in a modified buffer containing single stand binding protein (200 ng/ul) and spermidine (0.04 mM) to increase the sensitivity of amplification (52). DNA quality was assessed by ratio of A260/280 and A260/230 using a NanoDrop ND-1000 Spectrophotometer (Thermo Fischer Scientific, USA), and final DNA

concentrations were quantified with a PicoGreen method (53). DNAs were labeled and all hydridizations using the GeoChip 3.0 and image processing were conducted as previously described (54). Raw data from Imagene were analyzed using a GeoChip data analysis pipeline, where a signal to noise ratio of ≥1.5 was considered a positive signal (55).

Statistical Analysis. To visualize relationships between bacterial community composition and environmental variables, we ordinated community data (97% OTUs) using nonmetric multidimensional scaling (NMS), and Sorensen (Bray-Curtis) distance measures. NMS is an indirect ordination method that is well suited for non-normal data or data which occur along arbitrary or discontinuous scales. Detrended correspondence analysis (DCA) was used to examine the heterogeneity of the functional community composition. DCA is a commonly used ordination technique which evaluates the heterogeneity of community composition with functional gene data from the GeoChip (56). Kendall correlations were used to examine relationships between environmental variables and the NMS axes, and vectors were graphed to indicate both the direction and magnitude of the most strongly correlated variables in the ordination space.

Multiple response permutation procedure (MRPP) and Mantel tests were used to evaluate the relationship between microbial community structure or function and categorical environmental attributes such as successional stage and vegetation type. Relationships between individual OTUs or functional genes and environmental factors, such as successional stage, were tested using an indicator species analysis. This analysis can be used in association with the functional genes because the GeoChip produces both abundance and frequency data. Calculations were carried out using PC-Ord 5.0 (McCune and Mefford 1999).

RESULTS

Microbial community diversity and structure. Pyrosequencing resulted in a minimum of 735 sequences from each of the tree stands after sequence processing and normalization. These sequences grouped into between 213 and 791 OTUs at the 97% similarity level (Table 3-1). Forest stands dominated by white spruce (UP3A/WhiteSpruce1), aspen (UP1C), and birch (UP1B) had the highest bacterial species richness as measured by the (normalized) Chao 1 index. Black spruce-dominated stands had the lowest overall bacterial species richness, excluding BlackSpruce3 (TKN0022). The birch stand (UP1B) had the lowest number of OTUs obtained after normalization (Table 3-1, Figure S3-1).

The majority of 16S rRNA gene sequence retrieved from the four different vegetation types clustered with *Acidobacteria*, *Actinobacteria* and *Proteobacteria* – together these three phyla represented 83% to 97% of all sequences (Figure S3-2). In stands dominated by black spruce (TKN0001, TKN0012, TKN 0022), 40% of bacterial sequences clustered with the phylum *Proteobacteria* (Table 3-2). Of these, 17% of sequences in these stands belonged to the class *Gammaproteobacteria* and 5% clustered with *Bradyrhizobiaceae* (of the class *Alphaproteobacteria*). *Proteobacteria* also constituted a large portion (32-35%) of the soil community in white spruce (UP3A, UP3B, UP3C), aspen (UP1C) and birch (UP1B) dominated stands – with *Rhizobiales* (*Alphaproteobacteria*) and unclassified *Gammaproteobacteria* representing the majority of Proteobacterial sequences present in all stand types. In all forest stand types, taxa clustering with *Actinobacteria* constituted between 28-42% of the microbial community. In white spruce (UP3A, UP3B, UP3C) and birch (UP1B) stands, unclassified *Actinomycetales* and *Solirubrobacterales* dominated the *Actinobacterial* population, comprising 4-8% of the community. In the aspen-dominated stand (UP1C), sequences clustering with

Mycobacterium were the most abundant *Actinobacteria* (5% of the population). Sequences identified as *Acidobacteria* also constituted a large portion of the community in all soils tested: 25% of the population in black spruce stands, 13% of the population in white spruce stands, 13% of the birch associated sequences and 11% of the aspen associated community.

The relationship between bacterial community structure and aboveground vegetation was assessed using nonmetric multidimensional scaling (NMS) of community composition data using the Bray-Curtis Similarity measure. NMS ordination and multiple response permutation procedure (MRPP) analyses detected differences in microbial community structure across stand types (Figure 3-1, Table 3-3). In the MRPP tests, stand type had the largest effect size, with black spruce stands significantly different from both the early successional (birch and aspen stands) and the later successional white spruce stands. The differentiation in community structure across successional stages was reflected in the NMS ordination of sites (R^2=0.744, Stress = 0.1005; Figure 3-1). The lack of differentiation between the early (birch and aspen) and the late white spruce may be explained by the presence of white spruce in the birch sites. Superimposed on the NMS ordination are vectors indicating environmental drivers of the communities within ordination space. The bi-plots indicate that increasing soil C and decreasing pH are associated with communities in late-black spruce sites, while increasing cover of evergreen trees are associated with communities of the late-white spruce sites (Figure 3-1). Mantel tests were also performed to examine correlations between bacterial community structure and measured environmental and vegetation variables. The mantel test indicated that bacterial community structure was significantly and strongly correlated to vegetation cover ($p = 0.0009$, $r = 0.6490$). Bacterial community structure also correlated to measured environmental variables, however to a lesser extent than vegetation cover ($p = 0.0344$, $r = 0.3729$).

Indicator species analysis showed that there are diverse taxa that are indicators for each stand type. The classes *Actinobacteria* and *Acidobacteria* were present as indicators of each of the successional stages tested; early (aspen and birch dominated stands), late white spruce dominated stands, and late black spruce dominated stands. Twenty eight OTUs were significant indicators of early successional stage forest soils (indicator value = 100, $P < 0.038$) including members of the classes *Acidobacteria, Actinobacteria, Firmicutes, Alpha-*, *Beta-* and *Gamma-Proteobacteria* (Table 3-4). Sixteen OTUs were significant indicators of the mature, white spruce stands (indicator value = 100, $P = 0.034$), including members of the classes *Acidobacteria, Actinobacteria, Bacteroidetes*, and *Alpha-* and *Beta- Proteobacteria*. Soils from black spruce stands had two taxa which were significant indicators; *Calditerricola satsumensis* and *Acidobacterium capsulatum* (Table 3-4).

Microbial Community Functional Potential. Community functional gene diversity, richness and evenness were assessed for each of the tested soils based on microarray results (Table 3-1). Average Simpson's reciprocal diversity index (1/D) was highest in white spruce stands and lowest in the birch-dominated stands, with overall diversity decreasing from white spruce to black spruce, to aspen and birch. Similar results were observed with the Shannon-Weaver index. Evenness was comparable for all spruce and aspen sites, while it was lower in the birch stand (Table 3-1). White spruce and black spruce had the greatest number of unique genes (1031, 676 respectively). The birch and aspen stands had the highest proportion of overlapping genes between any two vegetation types, with 152.

Overall measured functional gene intensity (presence and amount of each gene detected) for genes measured on the GeoChip was highest in the white spruce 3 stand (UP3C), followed by the black spruce stands 1 and 2 (TKN 0001, TKN0022) (Figure S3-3a). The functional gene

intensity for genes associated with aromatic compound degradation and C cycling also followed this trend, with the highest gene intensities found in the white spruce 3 (UP3A), followed by the black spruce stands 1 and 2 (TKN 0001, TKN0022) (Figure S3-4a, Figure S3-5a). When examining the relative gene intensity (as a percentage of all genes detected within that soil type), genes associated with organic remediation represented the highest proportion of genes detected in all soils, except for early successional birch soils (UP1B), which were dominated by genes associated with metal resistance (Figure S3-3b). The second most abundant gene group in early succession birch (UP1B) soils were those associated with organic remediation, where in all other soils, genes associated with metal resistance were the second most abundant (Figure S3-4b). Gene associated with C and N cycling were the third and fourth most abundant gene groups in all soils measured (Figure S3-3b). White spruce 3 soils (UP3A) had the highest proportion of genes associated with C, phosphorous and sulfur cycling, energy processing (e.g. cytochrome, hydrogenase), and other genes (e.g. *gyrB*). The late successional stage black spruce (TKN0012) soils had highest proportion of N cycling genes (Figure S3-3b).

Within the category of "organic remediation genes", late successional stage soils had the highest proportion of aromatic carboxylic acid genes (white spruce 3, UP3C), BTEX and related aromatics (white spruce 2, UP3B), chlorinated aromatic genes (white spruce 3, UP3C), and nitroaromatic genes (white spruce 2, UP3B). Late successional stage black spruce forest soils had the highest proportion of "other" aromatic processing genes (e.g. catechol, *pcaG*, *nitA*, *pheA*, *proO*, *mdlD*, and *AmiE*; TKN0012; Table 3-4). Early successional soils dominated by aspen (UP1C) had the highest proportion of genes associated with degradation of heterocyclic and polycyclic compounds (Figure S3-4b).

The late successional stage, black spruce dominated soils had the highest proportion of C cycling genes from the categories C fixation (TKN0012), hemicellulose (TKN0001), pectin degradation genes (TKN 0022), and starch degradation genes (TKN0001). Late stage white spruce dominated soils had the highest proportion of cellulose (UP3C) and lignin (UP3B) degradation genes. Aspen dominated early succession soils (UP1C) had the highest proportion of genes associated with chitin degradation (Figure S3-5b).

The relationship between microbial community functional potential and the above ground vegetation with respect to measured environmental variables and successional stage was assessed using detrended correspondence analysis (DCA). Multiple response permutation procedure (MRPP) analyses and DCA ordinations detected differences in microbial community function between the white and black spruce stands and between the upland (UP) and lowland (TKN) sites (Figure 3-2, Table 3-2). In the MRPP tests, the late stage white and black spruce sites were significantly different from each other (Table 3-3). The microbial community functional potential also varied significantly between the sites located in the uplands (aspen, birch and white spruce) versus the lowland black spruce sites. The differentiation in community function across stand types and between young and mature successional stages was reflected in the DCA ordination of sites (3 axes; Figure 3-2). Superimposed on the DCA ordination are vectors indicating environmental correlates of the gene sets within ordination space. The environmental variables included in the DCA analysis indicate that soil chemistry and cover of trees and grasses are major factors influencing the functional differences between the forest types. These data suggest that the microbial functional potential associated with the different tree stands are unique to each of the stand types and strongly influenced by the underlying environmental chemistry. Mantel tests were also performed to test for correlations between microbial community

functional potential and measured environmental and vegetation variables. Interestingly, the mantel test indicated that functional potential was not significantly correlated to either measured environmental parameters ($p = 0.2056$, $r = 0.2842$) or the vegetation cover ($p = 0.7203$, $r = -0.0674$).

Indicator analysis showed that different genes associated with organic remediation were indicators for each stand type, with diverse functional genes present in each of the stand types (Table 3-5). Twenty two functional genes were significant indicators for soils from black spruce dominated forests (indicator value = 67.3-100, $P < 0.042$), including genes associated with C cycling (5 genes), metal resistance (4 genes), N cycling (3 genes), organic remediation (7 genes), phosphorus cycling (1 gene), and sulfur cycling (2 genes) (Table 3-5). Six functional genes were significant indicators of late successional white spruce soils, including genes associated with C cycling (2 genes), N cycling (2 genes), and organic compound remediation (2 genes). Early successional forest soils had two functional genes which were significant indicators; metal resistance gene *czcA*, and organic remediation gene *exaA*.

qPCR estimates of bacterial and fungal abundance. The abundance of bacteria and fungi, as determined using qPCR, did not show a consistent relationship with forest stand type in the tested DNA extract ($p > 0.05$, $R^2 < 0.002$). There was, however, a significant relationship between fungal and bacterial small subunit rRNA gene copies ($p = 0.01$, $R^2 = 0.4352$), with increasing bacterial gene copies corresponding with an increase in fungal gene copies.

The number of functional genes detected in soils as by the GeoChip did not correlate to the amount of bacterial 16S rRNA or fungal ITS1 gene copies in the environment (16S: $R^2 < 0.0004$, $p > 0.05$; ITS1 $R^2 = 0.301$, $p > 0.1$). This test was done in an effort to ensure that the observed functional genetic diversity is a measurement of the microbial community and not an

artifact or function of methodological constraints related to the relative abundance of bacterial or fungal DNA.

DISCUSSION

In this study, we sought to better understand the environmental drivers of microbial community structure and functional potential in boreal forest soils. Through characterization of bacterial community structure and analysis of microbial functional genes encoding key enzymes involved in major biogeochemical processes we found that vegetation type, associated with successional stage, appears to be the dominant factor driving the soil community. In addition, genes involved in aromatic compound degradation, including xenobiotic compounds, were present in each of the forest stand types at varying levels. Finally, vegetation also appears to have an effect on the number of genes associated with biogeochemical cycling in soil communities.

Microbial Community Structure. Through extensive sequencing of both bacterial and fungal communities, the combined efforts of this study and a previous study (24, 37) provide insight into the diverse microbial communities present in boreal forest soils. Recent studies using a variety of molecular or biochemical approaches have explored the distributional patterns exhibited by soil microbial communities and the biotic or abiotic factors driving those patterns. However, many of these previous studies have used techniques which do not permit detailed and comprehensive phylogenetic or taxonomic surveys of microbial communities across successional stages. Examples in the boreal forest include phospholipid fatty acid (PLFA) analyses (23, 57), bacterial and fungal clone libraries (58), and 16S rDNA pyrosequencing (59). Each of these studies has indicated that soil microbial communities within dominant vegetation types are more

similar to each other than to communities from different vegetation types or successional stages. However, none of these studies has comprehensively examined both the bacterial and fungal communities. Both the bacterial and fungal communities in the investigated soils have distinct populations associated with different stand types. In the fungal community only a few genera, such as *Cortinarius* (24) occur in the organic horizon of all major stand types. The composition of the bacterial community structure across successional stages was similar at the phylum level, with *Acidobacteria*, *Actinobacteria* and *Proteobacteria* representing the majority of sequences, but with differences in the composition within each phylum. In late stage successional black spruce sites which had the lowest pH, sequences grouping with *Acidobacteria* had the highest abundance. *Acidobacteria* is often associated with sites that have low pH, and has been previously found in boreal black spruce sites (4, 60).

The indicator species analysis showed that bacterial taxa belonging to the same class often have diverse preferences with respect to successional stage and plant community; however there are diverse taxa that are indicators for each forest stage (Table 3-4). This analysis highlighted the bacterial taxa whose variable abundance across samples significantly contributed to differences in bacterial community composition across successional stages, and is similar to fungal results (24, 37). All of the highly abundant classes were implicated in multiple stands, however none of the OTUs found as an indicator in one successional stage was significantly indicative of another successional stage. Overall, the abundant OTUs that were implicated as indicator species are numerically abundant and are species which are potentially important to soil functioning in different forest stand types. The functional abilities of indicator species warrants further investigation.

With the observed differences between microbial communities in different forest stand types successional stages and vegetation types, the question remains as to what is driving this difference. Ordination and mantel tests of the bacterial community indicated that percent cover of vegetation (i.e. lichen, evergreen tree, forbs and non-vegetation such as dead plant material) and, to a lesser extent, environmental factors (i.e. pH and C) were the main drivers of the separation between successional stages (Figure 3-1). MRPP tests are consistent with the mantel and NMS, in which the three successional stages or upland (aspen, birch and white spruce) versus the lowland black spruce sites are significant predictors for the bacterial community. These results suggest that the structure of the microbial community in the organic layer is correlated to the composition of the dominant vegetation which influences and is influenced by the soil environment.

Microbial Functional Potential. It is important to analyze microbial functional genes encoding key enzymes involved in major biogeochemical processes and contaminant degradation in an effort to understand what drives microbial functional potential. Many studies have investigated cellulose degradation (23), enzymatic activity (58), and C or N cycling (20, 61, 62) in the boreal forest, particularly in terms of understanding how biogeochemical cycles may change in response to global climate change. However, little is known regarding the functional communities associated with different successional stages in the Alaskan boreal forest. Through the use of the GeoChip functional gene microarray, we are able to gain a more comprehensive understanding of the microbial functional potential of soils, including biogeochemical cycling (i.e. C, N, P, and S), aromatic and xenobiotic compound degradation, and metal resistance.

The microarray data revealed overall similarities of functional potential within each successional stage. These relationships were observed through the use of ordination techniques

(Figure 3-2) and tested using MRPP, which showed that functional communities are statistically different only between the white and black spruce dominated soils (Table 3-3). The ordination also indicated that above ground vegetation (i.e. percent cover of deciduous trees & shrubs, evergreen trees & shrubs, graminoids, and lichens) and environmental parameters (i.e. soil CEC, pH, C) were strongly correlated to the different gene composition in each of the successional stages (Figure 3-2). The mantel test however, indicated that neither environmental chemistry nor vegetation were significant predictors of the functional community. The uncertainty in findings may be attributed to the sheer number of genes measured with the GeoChip functional gene microarray: 6165 total genes measured across eight different functional gene categories. Previously pH and soil moisture have been found to affect functional potential, measured using multi-substrate-induced respiration (multi-SIR) (61, 63). Measurement of soil enzyme activities among vegetation types has also been found to affect microbial cycling of C and N (58, 62). However, Priha et al. (23) using a culture based technique examining broad level microbial functioning (community level physiological profiling), failed to distinguish microbial community functioning from different boreal tree species. Important environmental measures such as ratios of C:N, lignin: N, or lignin:cellulose were not assessed in the current study, indicating that the lack of statistical relationship could be because an important factor was not included in the analysis (64).

Interestingly, in each of the investigated tree stands there were large proportions of genes associated with contaminant degradation (Figures S3-3, S3-4, and S3-5). Genes for contaminant degradation and organics remediation were determined to be indicators for each successional stage, indicating that they differ among forest stands community (Table 3-5). This suggests that as disturbance and the potential for disturbance, in the form of organic contamination, increases

in boreal regions the microbial communities may have the potential to acclimate and metabolize the contaminants. Although some contamination may be present at the site due to atmospheric deposition, the investigated sites are largely pristine. The questions arise as to what is driving the microbial community to produce genes associated with organics remediation and how will the community respond to disturbance. Recent work has suggested that plant secondary metabolites (PSM) may have an effect on the pollutant degradation capacity of microbial populations (35), and that different PSM may induce differential degradation of contaminants (65). The differential potential for degradation exhibited by the investigated vegetation indicates that certain vegetation types exhibit potential for future phytoremediation studies. Based on the genes observed, aspen and birch-dominated forest stands have the highest potential for degradation of chlorinated solvents, due to their association with *exaA*, a gene associated with dichloroethane degradation.

Overall, our data show that bacterial community structure and community function in the investigated boreal forest soils is dictated by dominant vegetation type and environmental condition. In addition, there is a high potential for the microbial degradation of xenobiotic contaminants, despite the lack of actual contamination in investigated soils. Future studies should use additional measures of activity, such as mRNA (metatranscriptomics) and functional assays, such as enzyme assays, to assess microbial community function. In addition, studies investigating the role of phytochemistry, including plant secondary metabolites, in driving microbial community function would be valuable to understanding the mechanisms underlying phytoremediation.

ACKNOWLEDGEMENTS

The project described was supported by Grant Number 5P20RR016466 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH) and a graduate fellowship from the Alaska Idea Network for Biomedical Research Excellence (INBRE), and EPSCoR NSF award #EPS-0701898 & the State of Alaska. MCL acknowledges the David L. Boren National Security Education Program (NSEP) Fellowship for support.

We thank Jakub Rídl and Miluse Hroudová of the Institute of Molecular Genetics, Czech Academy of Sciences for help with sequencing. We also thank Dr. Jizhong Zhou, Dr. Joy Van Nostrand and the Institute for Environmental Genomics at the University of Oklahoma for help with the GeoChip sample preparation and data processing. Finally, thanks to Dr. Todd O'Hara, Dr. Roger Ruess, and Dr. Teresa Hollingsworth at UAF for helpful editing, vegetation analysis and statistical advice.

REFERENCES

1. **Nannipieri P, Ascher J, Ceccherini MT, Landi L, Pietramellara G, Renella G**. 2003. Microbial diversity and soil functions. Eur. J. Soil Sci. **54**:655–670.

2. **Frossard A, Gerull L, Mutz M, Gessner MO**. 2011. Disconnect of microbial structure and function: enzyme activities and bacterial communities in nascent stream corridors. ISME J. 1–12.

3. **Monard C, Vandenkoornhuyse P, Le Bot B, Binet F**. 2011. Relationship between bacterial diversity and function under biotic control: the soil pesticide degraders as a case study. ISME J. **5**:1048–56.

4. **Lauber CL, Hamady M, Knight R, Fierer N**. 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl. Environ. Microbiol. **75**:5111–20.

5. **Rousk J, Bååth E, Brookes PC, Lauber CL, Lozupone C, Caporaso JG, Knight R, Fierer N**. 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. ISME J. **4**:1340–51.

6. **Brockett BFT, Prescott CE, Grayston SJ**. 2012. Soil moisture is the major factor influencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western Canada. Soil Biol. Biochem. **44**:9–20.

7. **Sagova-Mareckova M, Omelka M, Cermak L, Kamenik Z, Olsovska J, Hackl E, Kopecky J, Hadacek F**. 2011. Microbial communities show parallels at sites with distinct litter and soil characteristics. Appl. Environ. Microbiol. **77**:7560–7.

8. **Yergeau E, Sanschagrin S, Maynard C, St-Arnaud M, Greer CW**. 2013. Microbial expression profiles in the rhizosphere of willows depend on soil contamination. ISME J. 1–15.

9. **Xie J, He Z, Liu X, Liu X, Van Nostrand JD, Deng Y, Wu L, Zhou J, Qiu G**. 2011. GeoChip-based analysis of the functional gene diversity and metabolic potential of microbial communities in acid mine drainage. Appl. Environ. Microbiol. **77**:991–9.

10. **Shi W, Becker J, Bischoff M, Turco RF, Konopka AE**. 2002. Association of Microbial Community Composition and Activity with Lead, Chromium, and Hydrocarbon Contamination. Appl. Environ. Microbiol. **68**:3859–3866.

11. **Robertson SJ, Kennedy NM, Massicotte HB, Rutherford PM**. 2010. Enhanced biodegradation of petroleum hydrocarbons in the mycorrhizosphere of sub-boreal forest soils. Environ. Microbiol. Rep. **2**:587–593.

12. **Thomson BC, Ostle N, McNamara N, Bailey MJ, Whiteley AS, Griffiths RI**. 2010. Vegetation affects the relative abundances of dominant soil bacterial taxa and soil respiration rates in an upland grassland soil. Microb. Ecol. **59**:335–43.

13. **Hartmann A, Schmid M, Tuinen D Van, Berg G**. 2008. Plant-driven selection of microbes. Plant Soil **321**:235–257.

14. **Palmroth MRT, Münster U, Pichtel J, Puhakka JA**. 2005. Metabolic responses of microbiota to diesel fuel addition in vegetated soil. Biodegradation **16**:91–101.

15. **Kuske CR, Ticknor LO, Miller ME, Dunbar JM, Davis JA, Barns SM, Belnap J**. 2002. Comparison of soil bacterial communities in rhizospheres of three plant species and the interspaces in an arid grassland. Appl. Environ. Microbiol. **68**:1854–63.

16. **Clemmensen KE, Bahr A, Ovaskainen O, Dahlberg A, Ekblad A, Wallander H, Stenlid J, Finlay RD, Wardle D a., Lindahl BD**. 2013. Roots and Associated Fungi Drive Long-Term Carbon Sequestration in Boreal Forest. Science (80-.). **339**:1615–1618.

17. **McGuire A, Melillo J, Kicklighter D**. 1997. Equilibrium responses of global net primary production and carbon storage to doubled atmospheric carbon dioxide: Sensitivity to changes in vegetation nitrogen concentration. Glob. Biogeochem. … **11**:173–189.

18. **Kasischke ES, Verbyla DL, Rupp TS, McGuire a. D, Murphy K a., Jandt R, Barnes JL, Hoy EE, Duffy P a., Calef M, Turetsky MR**. 2010. Alaska's changing fire regime — implications for the vulnerability of its boreal forests. Can. J. For. Res. **40**:1313–1324.

19. **Allison SD, Gartner TB, Mack MC, McGuire K, Treseder K**. 2010. Nitrogen alters carbon dynamics during early succession in boreal forest. Soil Biol. Biochem. **42**:1157–1164.

20. **Kane ES, Vogel JG**. 2009. Patterns of Total Ecosystem Carbon Storage with Changes in Soil Temperature in Boreal Black Spruce Forests. Ecosystems **12**:322–335.

21. **Mack MC, Schuur E a G, Bret-Harte MS, Shaver GR, Chapin FS**. 2004. Ecosystem carbon storage in arctic tundra reduced by long-term nutrient fertilization. Nature **431**:440–3.

22. **Pennanen T, Liski J, Bååth E, Kitunen V, Uotila J, Westman CJ, Fritze H**. 1999. Structure of the Microbial Communities in Coniferous Forest Soils in Relation to Site Fertility and Stand Development Stage. Microb. Ecol. **38**:168–179.

23. **Priha O, Grayston SJ, Hiukka R, Pennanen T, Smolander a.** 2001. Microbial community structure and characteristics of the organic matter in soils under Pinus sylvestris, Picea abies and Betula pendula at two forest sites. Biol. Fertil. Soils **33**:17–24.

24. **Taylor DL, Herriott IC, Stone KE, McFarland JW, Booth MG, Leigh MB**. 2010. Structure and resilience of fungal communities in Alaskan boreal forest soils. Can. J. For. Res. **40**:1288–1301.

25. **Kuiper I, Lagendijk EL, Bloemberg G V, Lugtenberg BJJ**. 2004. Rhizoremediation: a beneficial plant-microbe interaction. Mol. Plant. Microbe. Interact. **17**:6–15.

26. **Leigh MB, Prouzová P, Macková M, Macek T, Nagle DP, Fletcher JS**. 2006. Polychlorinated biphenyl (PCB)-degrading bacteria associated with trees in a PCB-contaminated site. Appl. Environ. Microbiol. **72**:2331–42.

27. **Donnelly P, Hegde RS, Fletcher JS**. 1994. Growth of PCB-degrading bacteria on compounds from photosynthetic plants. Chemosphere **28**:981–988.

28. **Ohtsubo Y, Nagata Y, Kimbara K, Takagi M, Ohta A**. 2000. Expression of the bph genes involved in biphenyl/PCB degradation in Pseudomonas sp. KKS102 induced by the biphenyl degradation intermediate, 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid. Gene **256**:223–8.

29. **Singer A, Crowley DE, Thompson IP**. 2003. Secondary plant metabolites in phytoremediation and biotransformation. Trends Biotechnol. **21**:123–130.

30. **Bryant JP, Provenza FD, Pastor J, Reichardt PB, Clausen TP, du Toit JT, Thomas P**. 1991. Interactions Between Woody Plants and Browsing Mammals Mediated by Secondary Metabolites. Annu. Rev. Ecol. Syst. **22**:431–446.

31. **Stark S, Julkunen-Tiitto R, Holappa E, Mikkola K, Nikula A**. 2008. Concentrations of foliar quercetin in natural populations of white birch (Betula pubescens) increase with latitude. J. Chem. Ecol. **34**:1382–91.

32. **Slater H, Gouin T, Leigh MB**. 2011. Assessing the potential for rhizoremediation of PCB contaminated soils in northern regions using native tree species. Chemosphere **84**:199–206.

33. **Sugai SF, Schimel JP**. 1993. Decomposition and biomass incorporation of 14c-labeled glucose and phenolics in taiga forest floor: effect of substrate quality, successional state, and season. Soil Biol. Biochem. **25**:1379–1389.

34. **McFarlin K**. 2010. Rhizoremediation of diesel-contaminated soil using Salix alaxensis. University of Alaska Fairbanks.

35. **Leewis M-C, Reynolds CM, Leigh MB**. 2013. Long-term effects of nutrient addition and phytoremediation on diesel and crude oil contaminated soils in subarctic Alaska. Cold Reg. Sci. Technol. **96**:129–137.

36. **He Z, Deng Y, Van Nostrand JD, Tu Q, Xu M, Hemme CL, Li X, Wu L, Gentry TJ, Yin Y, Liebich J, Hazen TC, Zhou J**. 2010. GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. ISME J. **4**:1167–1179.

37. **Taylor DL, Hollingsworth TN, McFarland JW, Lennon NJ, Nusbaum C, Ruess RW**. 2014. A first comprehensive census of fungi in soil reveals both hyperdiversity and fine-scale niche partitioning. Ecol. Monogr. **84**:3–20.

38. **Hollingsworth TN, Walker MD, Chapin III FS, Parsons AL**. 2006. Scale-dependent environmental controls over species composition in Alaskan black spruce communities. Can. J. For. Res. **36**:1781–1796.

39. **Peech M**. 1982. Hydrogen Ion Activity, p. 920–921. *In* Methods of Soil Anaysis, Part 1., 2nd ed. Soil Science Society of America, Madison, WI.

40. **Jackson M.** 1958. Soil Chemical Analysis. Prentice Hall.

41. **Michaelson GJ, Ping CL, Mitchell GA**. 1987. Correlation of Mehlich 3, bray 1, and ammonium acetate extractable P, K, Ca, and Mg for Alaska agricultural soils 1. Commun. Soil Sci. Plant Anal. **18**:1003–1015.

42. **Holmgren GS, Juve RL, Geschwender RC**. 1977. A Mechanically Controlled Variable Rate Leaching Device. Soil Sci. Soc. Am. Journal1 **41**:1207–1208.

43. **Day PR**. 1965. Particle Fractionation and Particle-Size Analysis, Hydrometer Method, p. 562–566. *In* Methods of Soil Anaysis, Part 1., 2nd ed. Soil Science Society of America, Madison, WI.

44. **Engelbrektson A, Kunin V, Wrighton KC, Zvenigorodsky N, Chen F, Ochman H, Hugenholtz P**. 2010. Experimental factors affecting PCR-based estimates of microbial species richness and evenness. ISME J. **4**:1–6.

45. **Schloss PD, Gevers D, Westcott SL**. 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One **6**:e27310.

46. **He R, Wooller MJ, Pohlman JW, Quensen J, Tiedje JM, Leigh MB**. 2012. Shifts in identity and activity of methanotrophs in arctic lake sediments in response to temperature changes. Appl. Environ. Microbiol. **78**:4715–23.

47. **He R, Wooller MJ, Pohlman JW, Quensen J, Tiedje JM, Leigh MB**. 2012. Diversity of active aerobic methanotrophs along depth profiles of arctic and subarctic lake water column and sediments. ISME J. **6**:1937–48.

48. **Leigh MB, Pellizari VH, Uhlík O, Sutka R, Rodrigues J, Ostrom NE, Zhou J, Tiedje JM**. 2007. Biphenyl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). ISME J. **1**:134–48.

49. **Gardes M, Bruns TD**. 1993. ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. Mol. Ecol. **2**:113–118.

50. **Vilgalys R, Hester M**. 1990. Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several Cryptococcus species. J. Bacteriol. **172**:4238–46.

51. **Uhlik O, Jecna K, Mackova M, Vlcek C, Hroudova M, Demnerova K, Paces V, Macek T**. 2009. Biphenyl-metabolizing bacteria in the rhizosphere of horseradish and bulk soil contaminated by polychlorinated biphenyls as revealed by stable isotope probing. Appl. Environ. Microbiol. **75**:6471–7.

52. **Yergeau E, Bokhorst S, Kang S, Zhou J, Greer CW, Aerts R, Kowalchuk G a**. 2011. Shifts in soil microorganisms in response to warming are consistent across a range of Antarctic environments. ISME J. 1–11.

53. **He Z, Deng Y, Van Nostrand JD**. 2009. GeoChip 3.0: Further Development and Applications of Functional Gene Arrays (FGAs) for Analysis of Microbial CommunitiesProgram.

54. **He Z, Gentry TJ, Schadt CW, Wu L, Liebich J, Chong SC, Huang Z, Wu W, Gu B, Jardine P, Criddle C, Zhou J**. 2007. GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J. **1**:67–77.

55. **Van Nostrand JD, Wu W-M, Wu L, Deng Y, Carley J, Carroll S, He Z, Gu B, Luo J, Criddle CS, Watson DB, Jardine PM, Marsh TL, Tiedje JM, Hazen TC, Zhou J**. 2009. GeoChip-based analysis of functional microbial communities during the reoxidation of a bioreduced uranium-contaminated aquifer. Environ. Microbiol. **11**:2611–26.

56. **Jackson D, Somers K**. 1991. Putting things in order: the ups and downs of detrended correspondence analysis. Am. Nat. **137**:704–712.

57. **Merilä P, Malmivaara-Lämsä M, Spetz P, Stark S, Vierikko K, Derome J, Fritze H**. 2010. Soil organic matter quality as a link between microbial community structure and vegetation composition along a successional gradient in a boreal forest. Appl. Soil Ecol. **46**:259–267.

58. **Dimitriu P a., Prescott CE, Quideau S a., Grayston SJ**. 2010. Impact of reclamation of surface-mined boreal forest soils on microbial community composition and function. Soil Biol. Biochem. **42**:2289–2297.

59. **Roesch LFW, Fulthorpe RR, Riva A, Casella G, Hadwin AKM, Kent AD, Daroub SH, Camargo FAO, Farmerie WG, Triplett EW**. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. ISME J. **1**:283–90.

60. **Dimitriu P a, Grayston SJ**. 2010. Relationship between soil properties and patterns of bacterial beta-diversity across reclaimed and natural boreal forest soils. Microb. Ecol. **59**:563–73.

61. **Rousk J, Brookes PC, Glanville HC, Jones DL**. 2011. Lack of correlation between turnover of low-molecular-weight dissolved organic carbon and differences in microbial community composition or growth across a soil pH gradient. Appl. Environ. Microbiol. **77**:2791–5.

62. **Allison SD, Czimczik CI, Treseder KK**. 2008. Microbial activity and soil respiration under nitrogen addition in Alaskan boreal forest. Glob. Chang. Biol. **14**:1156–1168.

63. **Swallow M, Quideau S a.** 2013. Moisture effects on microbial communities in boreal forest floors are stand-dependent. Appl. Soil Ecol. **63**:120–126.

64. **Talbot JM, Treseder KK**. 2012. Interactions among lignin, cellulose, and nitrogen drive litter chemistry-decay relationships. Ecology **93**:345–54.

65. **Uhlik O, Musilova L, Ridl J, Hroudova M, Vlcek C, Koubek J, Holeckova M, Mackova M, Macek T**. 2012. Plant secondary metabolite-induced shifts in bacterial community structure and degradative ability in contaminated soil. Appl. Microbiol. Biotechnol.

FIGURES

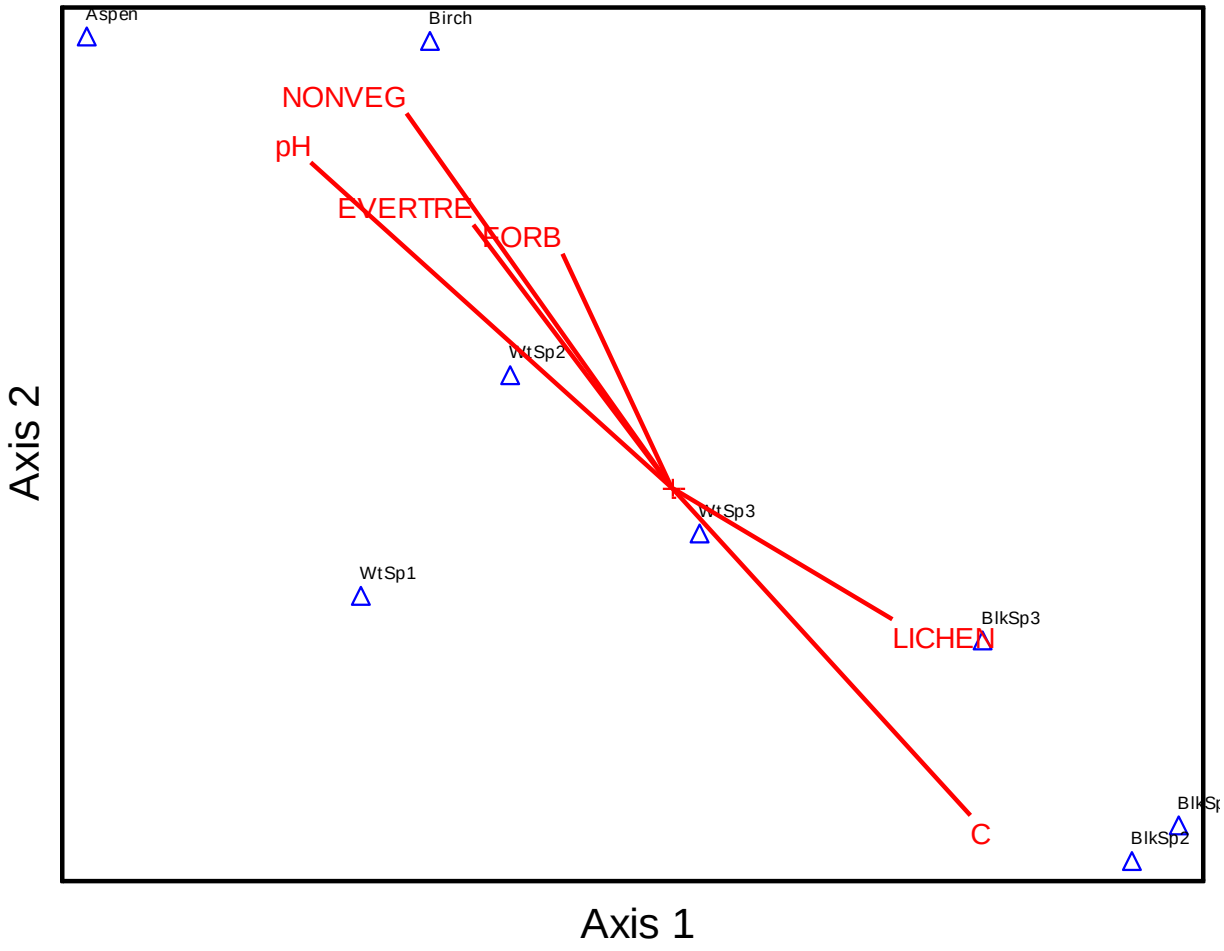

Figure 3-1: Non-metric multidimensional scaling (NMS) ordination of operational taxanomic units (OTUs) at 3% dissimilarity from eight samples in the BNZ LTER. Fitted vectors of selected ($r^2 > 0.5$) environmental variables showing significant correlation ($p < 0.05$) with the microbial community ordination spaces. The length of each vector shows the strength of the correlation in comparison with the other vectors. Environmental variables include percent cover of evergreen trees ("evertree"), grasses ("forb") or non-living matter ("nonveg", e.g. fallen trees, mushroom caps, burned material), soil C (C), and pH. Sample IDs have been shortened to White Spruce (W1, W2, W3), Black Spruce (B1, B2, B3), Aspen (A), and Birch (B).

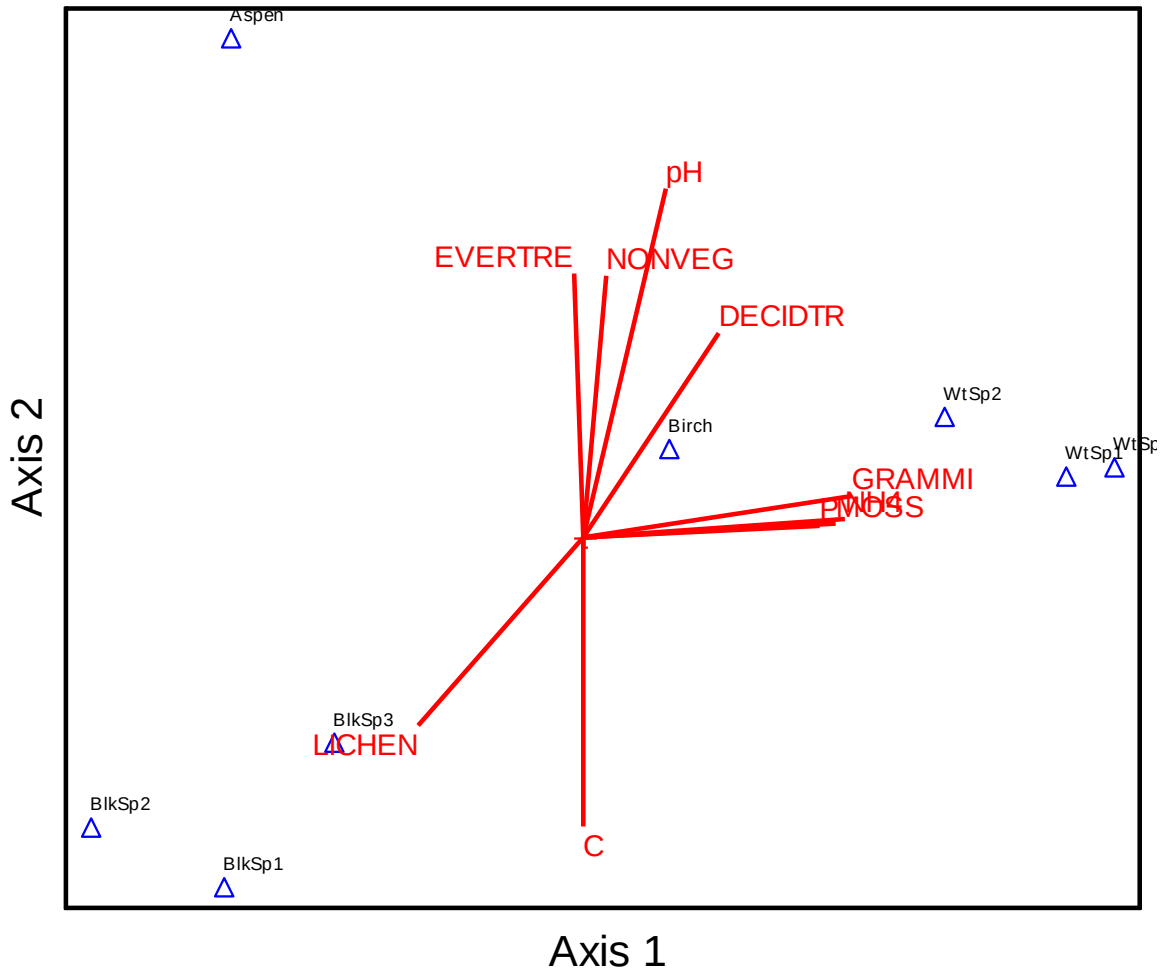

Figure 3-2: Detrended correspondence analysis (DCA) ordination of functional genes from eight samples in the BNZ LTER. All functional genes detected using the GeoChip 3.0 were used for the ordination. Fitted vectors of selected ($r^2 > 0.5$) environmental variables showing significant correlation ($p < 0.05$) with the microbial community ordination spaces. The length of each vector sows the strength of the correlation in comparison with the other vectors. Sample labels as in Figure 1.

TABLES

Table 3-1: Statistics for the phylogenetic sequence analysis and GeoChip analysis.

Site ID	Sample ID		Phylogenetic Statistics				Functional Gene Statistics			**Unique** & *Overlapped* Genes						
			Sequences after processing	Normalized Sequences	# OTUs (0.03)	Chao1	Total no. of genes detected	1/D (Inverse Simpson)	Shannon Weaver index	Birch	Aspen	Black Spruce1	Black Spruce2	Black Spruce3	White Spruce1	White Spruce2	White Spruce3
UP1B	Birch	Early Succession	12495	4328	4077	9285	296	170.88	5.43	**39**	-	-	-	-	-	-	-
UP1C	Aspen		9200	6006	2820	6339	571	380.95	6.15	*152*	**83**	-	-	-	-	-	-
TKN0001	BlackSpruce1	Late Succession	5325	5325	716	1703	1230	693.35	6.84	*205*	*375*	**260**	-	-	-	-	-
TKN0012	BlackSpruce2		6841	5440	714	1738	948	592.60	6.62	*146*	*262*	*432*	**309**		-	-	-
TKN0022	BlackSpruce3		5584	5429	1357	3293	340	210.74	5.61	*98*	*167*	*224*	*148*	**36**	-	-	-
UP3A	WhiteSpruce1		10273	6382	3181	7119	1889	1175.95	7.31	*230*	*422*	*801*	*509*	*249*	**740**	-	-
UP3B	WhiteSpruce2		7087	5863	2296	5539	83	58.99	4.25	*38*	*43*	*57*	*44*	*28*	*67*	**8**	
UP3C	WhiteSpruce3		7456	5580	2049	5045	808	453.38	6.43	*152*	*243*	*438*	*254*	*154*	*498*	*45*	**202**

Table 3-2: Relative abundance (% of total reads) and phylogenetic affiliations of all pyrosequencing reads from tested soils.

Phylogenetic Affiliations	Birch	Aspen	BlackSpruce1	BlackSpruce2	BlackSpruce3	WhiteSpruce1	WhiteSpruce2	WhiteSpruce3
Acidobacteria								
Acidobacteria								
Gp1	2.39 (299)	2.72 (250)	24.79 (1320)	26.31 (1800)	11.28 (630)	3.50 (360)	5.05 (358)	7.23 (539)
Gp2	0.33 (41)	0.10 (9)	0.81 (43)	1.05 (72)	2.60 (145)	0.97 (100)	0.47 (33)	1.02 (76)
Gp3	1.66 (207)	0.89 (82)	1.60 (85)	1.56 (107)	3.21 (179)	2.69 (276)	1.75 (124)	2.67 (199)
Gp6	3.50 (437)	2.39 (220)	0.08 (4)	-	0.75 (42)	2.21 (227)	1.96 (139)	1.30 (97)
Unclassified	-	0.01 (1)	0.02 (1)	-	0.04 (2)	0.10 (10)	0.03 (2)	0.05 (4)
Others	5.51 (689)	4.58 (421)	0.02 (1)	0.04 (3)	1.31 (73)	4.18 (429)	1.64 (116)	1.44 (107)
Actinobacteria								
Actinobacteria								
Acidimicrobiales	2.72 (340)	3.03 (279)	3.06 (163)	2.50 (171)	2.63 (147)	2.57 (264)	2.22 (157)	2.83 (211)
Actinomycetales	22.01 (2750)	27.37 (2518)	19.08 (1016)	20.63 (1411)	18.70 (1044)	17.88 (1837)	26.57 (1883)	23.22 (1731)
Solirubrobacterales	8.92 (1114)	9.25 (851)	4.83 (257)	2.85 (195)	6.39 (357)	10.39 (1067)	9.69 (687)	12.02 (896)
Thermoleophilales	0.22 (28)	0.13 (12)	-	-	-	0.18 (18)	0.10 (7)	0.13 (10)
Unclassified	3.55 (444)	2.70 (248)	0.66 (35)	0.48 (33)	1.56 (87)	4.71 (484)	2.85 (202)	1.97 (147)
Proteobacteria								
Alphaproteobacteria								
Rhizobiales	7.49 (936)	10.50 (966)	9.78 (521)	12.02 (822)	10.82 (604)	10.13 (1041)	10.86 (770)	11.43 (852)
Rhodospirillales	1.47 (184)	1.22 (112)	3.85 (205)	5.17 (354)	3.15 (176)	2.65 (272)	1.95 (138)	1.88 (140)
Others	1.90 (238)	2.29 (211)	0.83 (44)	0.73 (50)	1.84 (103)	2.02 (208)	2.34 (166)	2.27 (169)
Unclassified	0.35 (44)	0.35 (32)	0.45 (24)	0.41 (28)	0.54 (30)	0.54 (55)	0.32 (23)	0.54 (40)
Betaproteobacteria								
Burkholderiales	4.52 (565)	4.93 (454)	1.90 (101)	1.70 (116)	2.79 (156)	2.55 (262)	2.58 (183)	2.49 (186)
Others	0.18 (22)	0.12 (11)	-	-	0.07 (4)	0.25 (26)	0.06 (4)	0.07 (5)
Unclassified	1.41 (176)	0.99 (91)	0.13 (7)	0.10 (7)	0.47 (26)	1.41 (145)	0.73 (52)	0.38 (28)
Deltaproteobacteria								
Myxococcales	3.15 (394)	1.98 (182)	0.66 (35)	0.37 (25)	2.45 (137)	3.60 (370)	3.06 (217)	2.49 (186)
Others	0.26 (33)	0.47 (43)	0.06 (3)	0.01 (1)	0.09 (5)	0.25 (26)	0.14 (10)	0.05 (4)
Unclassified	1.14 (142)	0.72 (66)	0.41 (22)	0.26 (18)	0.98 (55)	1.12 (115)	0.68 (48)	0.75 (56)
Gammaproteobacteria								
Legionellales	0.43 (54)	0.46 (42)	1.16 (62)	0.32 (22)	0.32 (18)	0.65 (67)	0.38 (27)	0.51 (38)
Xanthomonadales	2.54 (317)	3.86 (355)	0.51 (27)	0.61 (42)	1.59 (89)	1.46 (150)	3.08 (218)	2.75 (205)
Others	1.03 (129)	0.68 (63)	0.13 (7)	0.09 (6)	0.32 (18)	0.86 (88)	0.87 (62)	0.50 (37)
Unclassified	3.36 (420)	3.37 (310)	19.70 (1049)	18.90 (1293)	11.14 (622)	3.77 (387)	5.55 (393)	6.55 (488)
Unclassified	2.76 (345)	2.63 (242)	0.96 (51)	0.66 (45)	2.17 (121)	3.66 (376)	3.02 (214)	2.52 (188)
Others	11.54 (1442)	8.26 (760)	3.25 (173)	2.43 (166)	7.61 (425)	8.58 (881)	8.47 (600)	6.93 (517)
Unclassified	5.64 (705)	4.01 (369)	1.30 (69)	0.79 (54)	5.18 (289)	7.13 (732)	3.58 (254)	4.02 (300)
Total	100 (12495)	100 (9200)	100 (5325)	100 (6841)	100 (5584)	100 (10273)	100 (7087)	100 (7456)

Table 3-3: Multiple-response permutation procedure tests for differentiation of bacterial community structure and total microbial community function.

	Community Structure		Community Function	
Factor	Effect size, A	Significance, P	Effect size, A	Significance, P
Successional stage	**0.2464139**	**0.01105106**	0.07405484	0.08896873
Early vs. Black Spruce	**0.27300627**	**0.0000001**	0.03192469	-
Early vs. White Spruce	0.0882196	-	0.02056518	-
Black Spruce vs. White Spruce	**0.23503459**	**0.0267207**	**0.10537371**	**0.03813113**
Aspect (Upland vs Lowland)	**0.19804364**	**0.00465736**	**0.06085169**	**0.04692471**

Table 3-4: Indicator species analysis showing successional stage (habitat) preferences for selected taxa. Significant indicator values indicate differentiation between successional stages.

Successional Stage	Phylum	Class	Closest RDP SeqMatch Type Strain (Similarity score)	Accession No.	Taxon	Observed Indicator Value	P
Black Spruce							
	Acidobacteria						
		Acidobacteria_Gp					
			Granulicella paludicola (1)	AM887758	Otu0009	58.4	0.024
			Acidobacterium capsulatum (0.746)	CP001472	Otu0018	100	0.032
			Edaphobacter modestus (0.855)	DQ528760	Otu0066	59.3	0.0212
	Actinobacteria						
		Actinobacteria					
			Mycobacterium nebraskense (0.895)	AY368456	Otu0004	98	0.0402
			Mycobacterium shimoidei (0.903)	AJ005005	Otu0005	91.2	0.004
			Solirubrobacter soli (0.752)	AB245334	Otu0007	63.1	0.007
			Rarobacter incanus (0.535)	AB056129	Otu0043	94.7	0.02
			Streptacidiphilus carbonis (1)	DSM 41754/AF074412	Otu0067	74.1	0.0222
			Luedemannella helvata (0.571)	AB236957	Otu0080	75	0.004
			Patulibacter americanus (0.613)	AJ871306	Otu0170	83.3	0.0444
			Calditerricola satsumensis '(0.454)	YMO81/AB250968	Otu0259	100	0.032
	Proteobacteria						
		Alphaproteobacteria					
			Pleomorphomonas oryzae (0.682)	AB159680	Otu0006	71.3	0.0288
			Methylocystis heyeri (0.733)	AM283543	Otu0138	71.4	0.0334
		Betaproteobacteria					
			Burkholderia sordidicola (0.946)	AF512827	Otu0017	94.1	0.0196
			Burkholderia sediminicola (0.941)	EU035613	Otu0077	91.7	0.0402
		Gammaproteobacteria					
			Marinobacter daepoensis (0.548)	AY517633	Otu0001	70	0.004
			Dokdonella fugitiva (0.752)	AJ969432	Otu0071	85.7	0.0104
		unclassified					
			Methylosinus sporium (0.563)	Y18946	Otu0011	96.8	0.0196
			Steroidobacter denitrificans (0.517)	EF605262	Otu0068	71.4	0.0392
	unclassified						
			Marinimicrobium agarilyticum (0.5)	AY839870	Otu0101	90.9	0.0196
Early							
	Acidobacteria						
		Acidobacteria_Gp					
			Acidobacterium capsulatum (0.724)	CP001472	Otu0167	100	0.0386
			Acidobacterium capsulatum (0.515)	CP001472	Otu0226	100	0.0386
			Geopsychrobacter electrodiphilus (0.435)	AY187303	Otu0310	100	0.0386
			Acidobacterium capsulatum (0.562)	CP001472	Otu0337	100	0.0386
	Actinobacteria						
		Actinobacteria					
			Solirubrobacter soli (0.691)	AB245334	Otu0094	100	0.0386
			Nocardioides dubius (0.727)	AY928902	Otu0151	100	0.0386
			Solirubrobacter soli (0.606)	AB245334	Otu0201	100	0.0386
			Solirubrobacter soli (0.806)	AB245334	Otu0237	100	0.0386
			Streptacidiphilus albus (0.966)	DSM 41753/AF074415	Otu0238	100	0.0386
			Sporichthya polymorpha (0.667)	AB025317	Otu0260	100	0.0386
			Cryptosporangium minutisporangium (0.736)	AB037007	Otu0294	100	0.0386
			Rarobacter incanus (0.5)	AB056129	Otu0301	100	0.0386
			Solirubrobacter soli (0.578)	AB245334	Otu0346	100	0.0386
			Aciditerrimonas ferrireducens (0.55)	AB517669	Otu0353	100	0.0386
			Mycobacterium wolinskyi (0.861)	AY457083	Otu0400	100	0.0386
			Streptomyces chryseus (0.919)	CSSP540/AY999787	Otu0423	100	0.0386
			Nocardioides dilutus (0.835)	EF466121	Otu0428	100	0.0386

Table 3-4 continued:

Successional Stage	Phylum	Class	Closest RDP SeqMatch Type Strain (Similarity score)	Accession No.	Taxon	Observed Indicator Value	P
	Firmicutes						
		Bacilli					
			Bacillus simplex (1)	AJ439078	Otu0332	100	0.0386
	Proteobacteria						
		Alphaproteobacteria					
			Acidiphilium angustum (0.757)	D30772	Otu0127	100	0.0386
			Acidisphaera rubrifaciens (0.73)	D86512	Otu0180	100	0.0386
			Magnetospirillum gryphiswaldense (0.696)	Y10109	Otu0385	100	0.0386
		Betaproteobacteria					
			Burkholderia sordidicola (0.89)	AF512827	Otu0242	100	0.0386
			Rugamonas rubra (0.824)	HM038005	Otu0323	100	0.0386
		Gammaproteobacteria					
			Steroidobacter denitrificans (0.733)	EF605262	Otu0254	100	0.0386
	unclassified						
			Cryptosporangium japonicum (0.433)	D85466	Otu0128	100	0.0386
			Catenulispora rubra (0.467)	AB180773	Otu0366	100	0.0386
			Cryptosporangium japonicum (0.49)	D85466/YU 636-3	Otu0442	100	0.0386
			Nocardia brasiliensis (0.444)	AF430038	Otu0473	100	0.0386
White Spruce							
	Acidobacteria						
		Acidobacteria_Gp3					
			Bryobacter aggregatus (0.556)	AM162405	Otu0190	100	0.034
		Acidobacteria_Gp4					
			Enterococcus pallens (0.4)	DQ411812	Otu0222	100	0.034
	Actinobacteria						
		Actinobacteria					
			Conexibacter woesei (0.746)	CP001854	Otu0026	79.1	0.0436
			Nocardioides halotolerans (0.749)	EF466122	Otu0076	84.2	0.034
			Blastococcus saxobsidens (0.843)	FN600641/BC444	Otu0111	100	0.034
			Solirubrobacter soli (0.782)	AB245334	Otu0252	100	0.034
			Streptomyces prunicolor (0.891)	DQ026659	Otu0309	100	0.034
			Humicoccus flavidus (0.713)	DQ321750	Otu0340	100	0.034
			Ferrimicrobium acidiphilum (0.463)	AF251436	Otu0354	100	0.034
			Streptomyces blastmyceticus (0.907)	CSSP556/AY999802	Otu0446	100	0.034
			Herbiconiux ginsengi (0.753)	DQ473536	Otu0483	100	0.034
			Solirubrobacter soli (0.677)	AB245334	Otu0518	100	0.034
	Bacteroidetes						
		Sphingobacteria					
			Chitinophaga ginsengisegetis (0.656)	AB264798	Otu0110	100	0.034
	Proteobacteria						
		Alphaproteobacteria					
			Agromonas oligotrophica (0.718)	D78366	Otu0052	92.3	0.02
			Kaistia adipata (0.629)	AY039817	Otu0164	100	0.034
			Caulobacter fusiformis (0.822)	AJ227759	Otu0304	100	0.034
			Afipia birgiae (0.855)	AF288304	Otu0378	100	0.034
		Betaproteobacteria					
			Burkholderia sordidicola (0.949)	AF512827	Otu0049	77.4	0.0226
			Burkholderia endofungorum (0.599)	AM420302	Otu0206	100	0.034
			Chitiniphilus shinanonensis (0.558)	AB453176	Otu0348	100	0.034
		Gammaproteobacteria					
			Kushneria avicenniae (0.587)	DQ888315	Otu0030	75	0.0272
			Dyella koreensis (0.887)	AY884571	Otu0095	87.5	0.0426
			Marinobacter excellens (0.525)	AY180101	Otu0172	83.3	0.038

Table 3-5: Indicator species analysis showing successional stage (vegetation) preferences for selected functional genes.

Successional Stage	Gene category	Compound	Gene name	Functional Gene	Observed Indicator Value	P
Black Spruce						
	Carbon cycling					
		Hemicellulose	ara	Func2432	100	0.031
		Carbon fixation	pcc	Func1216	89.9	0.042
		Carbon fixation	pcc	Func1909	89.7	0.019
		Lignin	phenol_oxidase	Func1487	88.3	0.019
		Hemicellulose	xylA	Func2386	100	0.031
	Nitrogen					
		Denitrification	narG	Func0510	85.1	0.018
		Denitrification	narG	Func1113	81.9	0.04
		Denitrification	nosZ	Func0603	73.6	0.009
	Organic Remediation					
		Alkanes (e.g. Octane)	alkK	Func2636	100	0.031
		Aromatic carboxylic acid (e.g. Benzoate)	benAB	Func0372	80.8	0.009
		Herbicides (e.g. Glyphosphate)	mauAB	Func2202	67.3	0.043
		Aromatic carboxylic acid (e.g. Salicylate)	nagG	Func0531	72.6	0.009
		Aromatic Compounds (e.g. Benzonitrile)	nitA	Func0945	92.9	0.019
		Herbicides (e.g. Glyphosphate, Formaldehyde)	phn	Func2592	100	0.031
		Aromatic carboxylic acid (e.g. Benzoate)	pimF	Func1679	87.5	0.019
	Metal Resistance					
		Arsenic	ArsC	Func1537	76.9	0.042
		Chromium	ChrA	Func0946	91.8	0.019
		Chromium	ChrA	Func0217	72.6	0.019
		Copper	CopA	Func2488	100	0.031
	Phosphorus					
		Phosphorus utilization	ppx	Func2056	95.7	0.019
	Sulphur					
		Sulfite Reductase	dsrA	Func2274	71.4	0.039
		Sulfite Reductase	dsrB	Func2087	93.6	0.019
Early						
	Organic Remediation					
		Chlorinated solvents (e.g. Dichloroethane)	exaA	Func1948	82.6	0.045
	Metal Resistance					
		Cadmium,Cobalt,Zinc	czcA	Func2670	100	0.031
White Spruce						
	Carbon cycling					
		Chitin	exochitinase	Func0129	100	0.036
		Hemicellulose	xylanase	Func0549	95.3	0.042
	Organic Remediation					
		BTEX and related aromatics (e.g. Ethylbenzene)	akbF	Func0623	100	0.036
		Aromatic carboxylic acid (e.g. Mandelate)	mdlA	Func0017	100	0.036
	Nitrogen					
		Denitrification	nirK	Func0378	97.3	0.021
		Denitrification	nirK	Func0383	90.5	0.042

SUPPLEMENTAL FIGURES

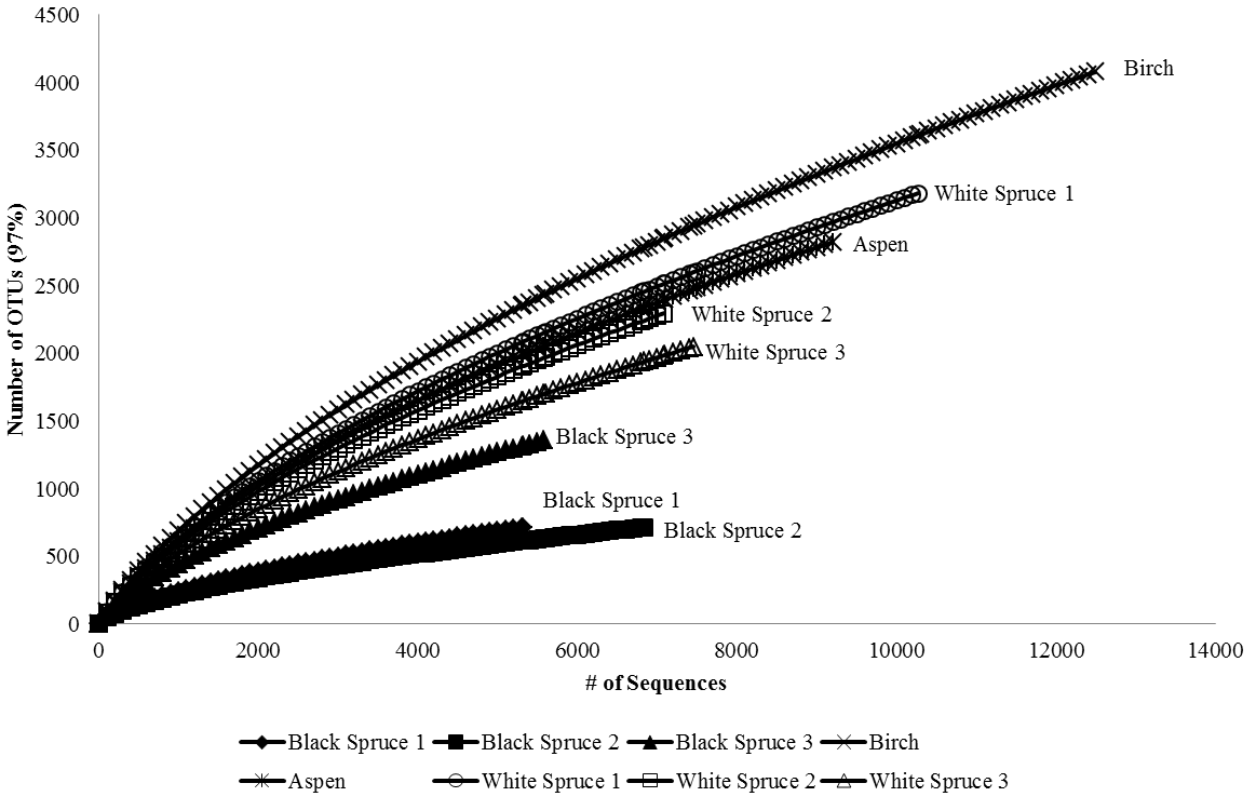

Figure S3-1: Rarefaction curves for the non-normalized numbers of pyrosequencing reads and operational taxanomic units (OTUs) at 3% dissimilarity.

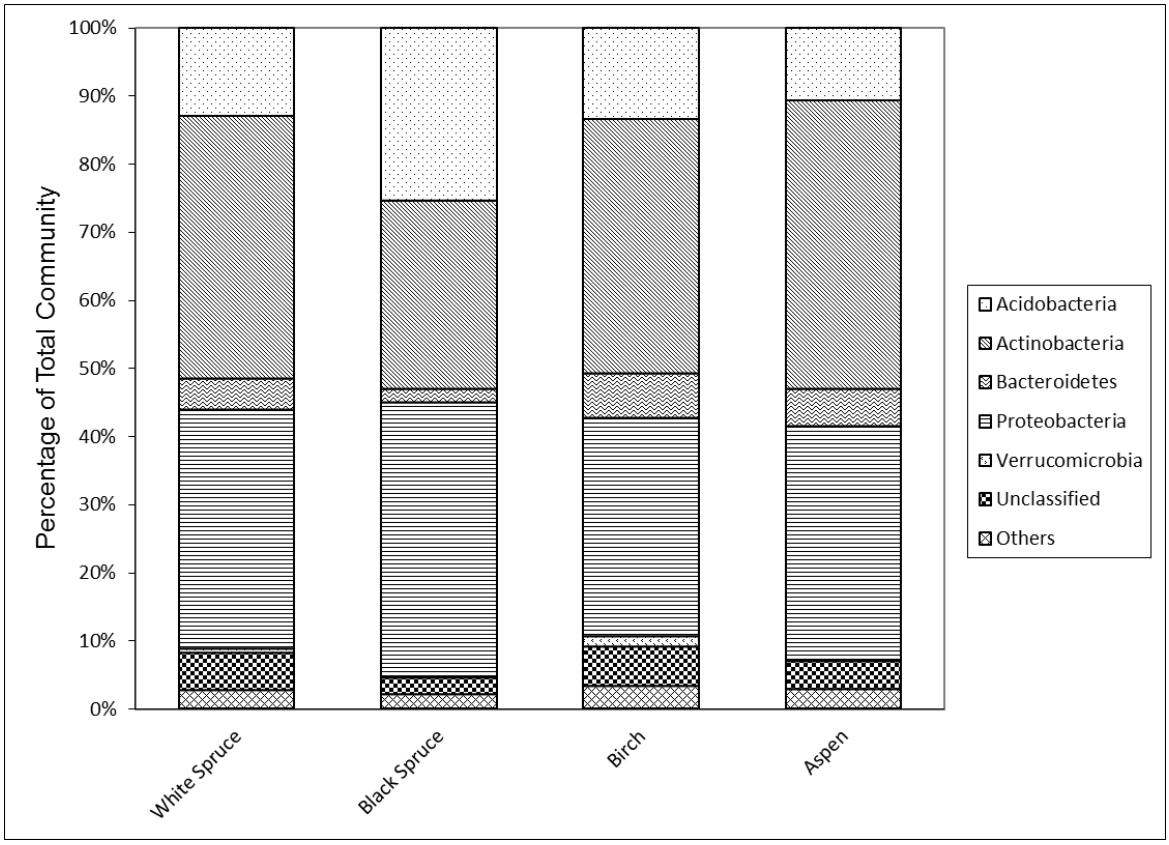

Figure S3-2: Soil bacterial communities as identified by 16S rRNA pyrosequencing represented by phylum (different textures) and abundance (relative size).

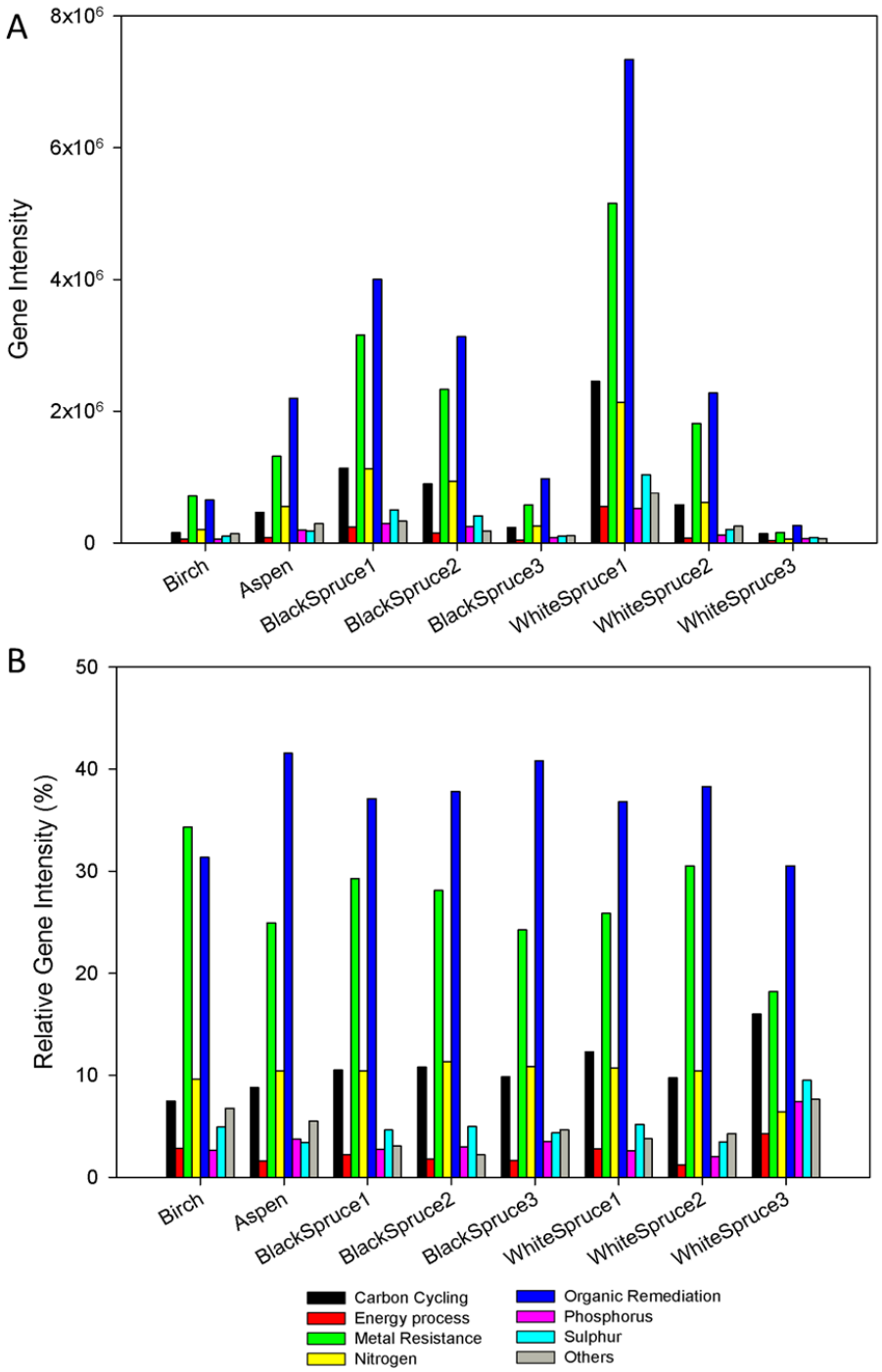

Figure S3-3: The total gene intensity (A) and relative abundance (B) of all detected genes from each of the vegetation types.

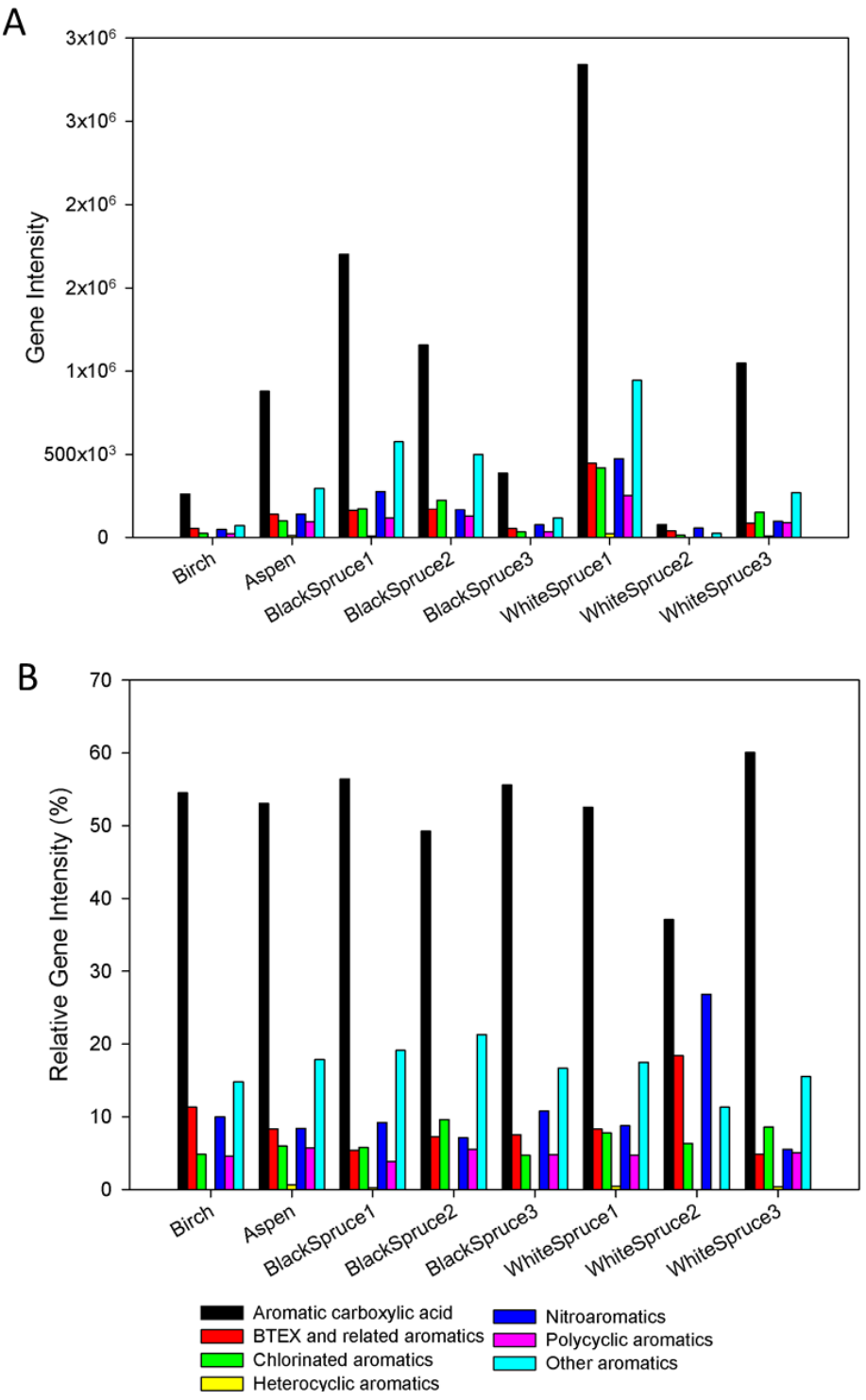

Figure S3-4: The total gene intensity (A) and relative abundance (B) of detected aromatic compound degradation genes from each of the vegetation types.

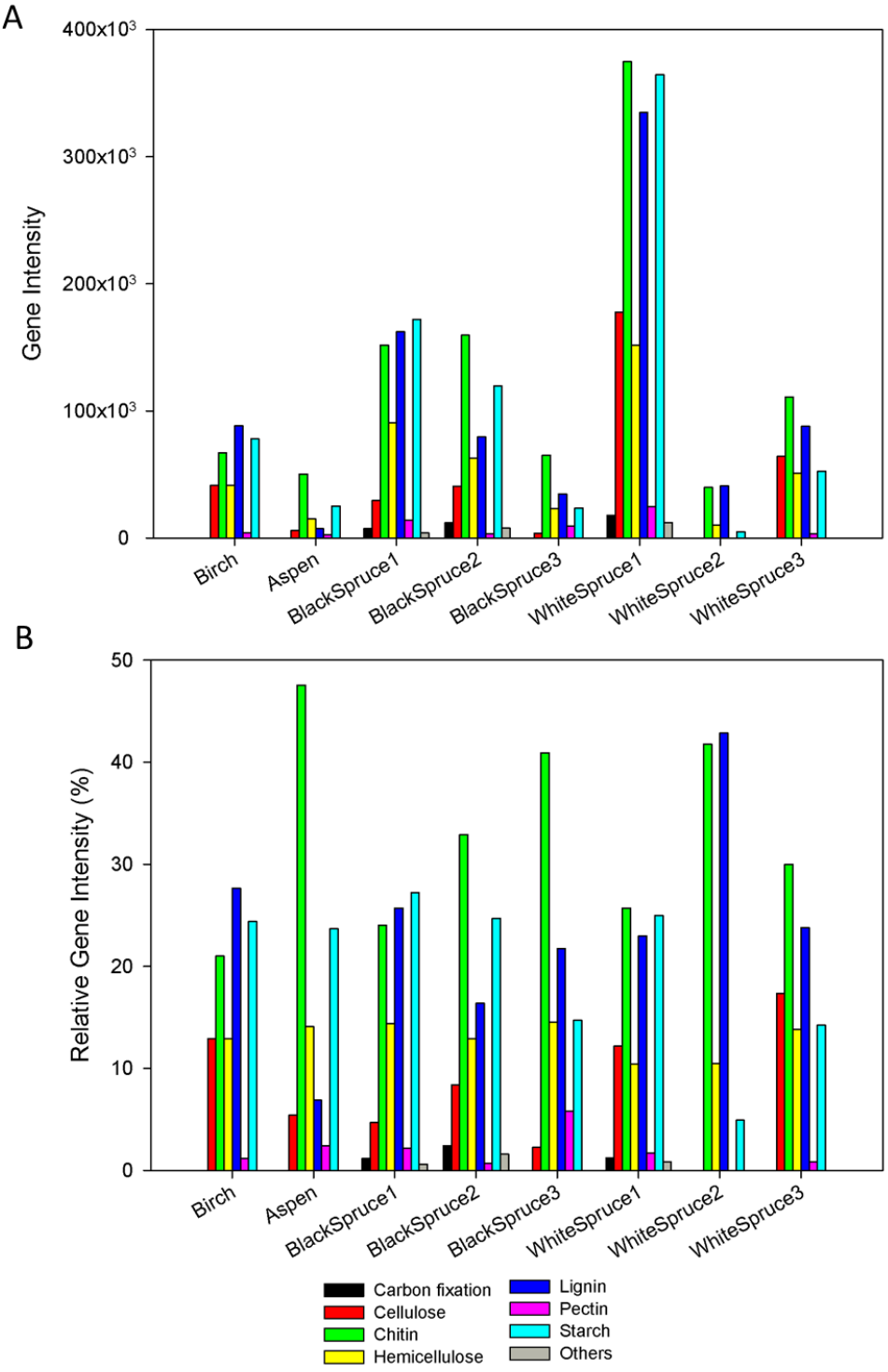

Figure S3-5: The total gene intensity (A) and relative abundance (B) of detected genes associated with C cycling from each of the vegetation types.

Table S3-1: Vegetation present at each site as represented by the type code. Numbers are Braun-Blanquét cover abundance scale, see Hollingsworth et al., 2006 for collection methods and further information (38).

Site ID	Sample ID		Deciduous Shrub	Deciduous Tree	Evergreen Shrub	Evergreen Tree	Forb	Gramminoids	Lichen	Moss	NonVeg*
UP1B	Birch	Early Succession	66	22	0	7	87	19	47	27	85
UP1C	Aspen		70	22	0	12	83	10	24	32	87
TKN0001	BlackSpruce1	Late Succession	13	1	0	8	6	2	16	27	0
TKN0012	BlackSpruce2		16	1	2	7	4	2	25	23	0
TKN0022	BlackSpruce3		15	0	0	8	12	3	15	30	0
UP3A	WhiteSpruce1		18	8	0	2	40	8	16	45	38
UP3B	WhiteSpruce2		34	7	0	4	40	5	9	38	42
UP3C	WhiteSpruce3		39	9	2	5	48	9	11	38	46

* NonVeg represents non-vegetation cover in the site such as standing dead wood, fungal fruitng bodies, rocks, and dead organic materials.

Table S3-2: Environmental and soil parameters for each of the investigated sites

Site ID	Sample ID		Vegetation Description	GPS coordinates		pH (1:1)	Ammonia	Nitrates	P (ppm)	K (ppm)	CEC (meq/100)	C (%)
				Easting	Northing							
UP1B	Birch	Early Succession	Young Paper Birch/Shrub/Herbaceous Stand	64.73463679	-148.3000982	5.87	42	<1	237	577	66.85	54.22
UP1C	Aspen		Young Aspen Forest	64.73463679	-148.3000982	6.50	41	<1	231	690	43.35	40.05
TKN0001	BlackSpruce1	Late Succession	black spruce	64.76572442	-148.295527	3.60	59	2	273	990	64.16	85.61
TKN0012	BlackSpruce2		black spruce	65.16721667	-147.8941333	3.67	44	<1	220	941	73.24	83.26
TKN0022	BlackSpruce3		black spruce	64.997	-147.65305	3.82	38	<1	226	1174	68.97	84.08
UP3A	WhiteSpruce1		Mature white spruce stand on a south-facing slope	64.76742521	-148.2795018	5.22	157	2	325	1096	74.59	71.00
UP3B	WhiteSpruce2		Mature white spruce stand on a south-facing slope	64.76051099	-148.2508148	5.22	256	8	391	1034	72.33	78.01
UP3C	WhiteSpruce3		Closed White Spruce Forest	64.72538077	-148.3306337	5.57	197	4	318	1108	86.64	76.03

Chapter 4

Long-term Effects of Nutrient Addition and Phytoremediation on Diesel and Crude Oil Contaminated Soils in subarctic Alaska [1]

ABSTRACT

Phytoremediation is a potentially inexpensive method of detoxifying contaminated soils using plants and associated soil microorganisms. The remote locations and cold climate of Alaska provide unique challenges associated with phytoremediation such as finding effective plant species that can achieve successful site clean-up despite the extreme environmental conditions and with minimal site management. A long-term assessment of phytoremediation was performed which capitalized on a study established in Fairbanks in 1995. The original study sought to determine how the introduction of plants (*Festuca rubra, Lolium multiflorum*), nutrients (fertilizer), or their combination would affect degradation of petroleum hydrocarbon (TPH) contaminated soils (crude oil or diesel) over time. Within the year following initial treatments, the plots subjected to both planting and/or fertilization showed greater overall decreases in TPH concentrations in both the diesel and crude oil contaminated soils relative to untreated plots. We re-examined this field site after 15 years with no active site management to assess the long-term effects of phytoremediation on colonization by native and non-native plants, their rhizosphere microbial communities and on petroleum removal from soil. Native and non-native vegetation had extensively colonized the site, with more abundant vegetation found on the diesel contaminated soils than the more nutrient-poor, more coarse, and acidic crude oil contaminated soils. TPH concentrations achieved regulatory clean up levels in all treatment groups, with lower TPH concentrations

[1] Leewis M-C, Reynolds CM, Leigh MB. 2013. Long-term effects of nutrient addition and phytoremediation on diesel and crude oil contaminated soils in subarctic Alaska. Cold Reg. Sci. Technol. 96:129–137.

correlating with higher amounts of woody vegetation (trees & shrubs). In addition, original treatment type has affected vegetation recruitment to each plot with woody vegetation and more native plants in unfertilized plots. Bacterial community structure also varies according to the originally applied treatments. This study suggests that initial treatment with native tree species in combination with grasses could be an effective means for phytoremediating petroleum contaminated soils and promoting ecological recovery in cold regions.

INTRODUCTION

At high latitudes, petroleum products are often the primary source of fuel in both urban and rural environments. Petroleum hydrocarbons (PHCs) can be introduced into the environment through natural seepage and surface spills such as leaking pipelines and storage tanks. There are more than 500 formerly used defense sites owned by the Department of Defense in Alaska with approximately 1000 individual areas of soil contamination, many contaminated with petroleum (Reynolds and Koenen, 1997; Yergeau et al., 2012). Although the need to remediate contaminated sites is not unique to high latitudes, remediation in these regions is often confounded by a lack of infrastructure, reduced degradation rates, and high expense associated with traditional remediation approaches. In Alaska, the human health risk of contaminated sites is also exacerbated by the fact that many rural communities are located off road systems and near contaminated sites, and these communities rely heavily upon subsistence food harvesting, which can increase the exposure of inhabitants to contaminants. Phytoremediation may offer one solution to feasibly reduce contaminant levels and human health risks in these communities.

There are currently two main remediation strategies for PHC contaminated sites; physical-chemical remediation by contaminant removal and bioremediation through the use of microorganisms and plants (Uhlik et al., 2009). Physical-chemical remediation methods are the conventional approach to remediation of contaminated soils; however, such methods can be extremely costly, involving removal and incineration or transfer to hazardous waste landfills. In remote regions, such remediation methods are often not available or are prohibitively expensive due to the lack of roads and other infrastructure (Slater et al., 2011). In addition to the expense, removal of contaminated soils by physical means can also leave much of the contaminant behind in the form of loose soil. Alternative methods of

contaminant abatement can include landfarming, biopiles and other forms of bioremediation such as phytoremediation (Palmroth et al., 2002; Paudyn et al., 2008).

Landfarming is common in cold regions. In landfarms, contaminated soils are spread over large areas and can be amended with nutrients, water, and surfactants in an effort to reduce the limitations associated with microbial degradative processes (Filler et al., 2006; Sanscartier et al., 2009). Periodic tilling may be performed to aerate soils and promote volatilization of contaminants. In biopile systems, contaminated soils are amended with nutrients and water then mounded over an aeration piping system. Soils can also be amended with organic material as a bulking agent to increase soil porosity and supply carbon. Aeration systems are varied and can be active or passive, and also could incorporate a heating system to increase soil temperature and therefore extend degradation treatment season (Sanscartier et al., 2009).

Both landfarming and biopiles are relatively low-cost and simple techniques and their use is wide spread across northern latitudes. However, these *ex-situ* strategies for contaminant abatement require the input of expensive elements such as heavy equipment for tillage, maintenance of aeration systems or reapplication of fertilizer multiple times a year. In remote regions with little or no road access, these strategies quickly become unreasonably expensive due to the cost of materials and travel to the site. Depending on the specific site management methods applied, these techniques may also leave contaminated soils exposed in stockpiles located at or near the community landfill. In addition, both landfarming and biopiles can disturb fragile ecosystems and can cause unnecessary destruction of the natural landscape.

Plant assisted bioremediation, or phytoremediation, is potentially an inexpensive and effective alternative method for cleanup of soil contaminants (Singer et al., 2003a; Wenzel, 2008). Phytoremediation is an attractive approach to bioremediation because it does not necessarily involve large soil disruptions and can result in intact, biologically active soils

with minimal financial input (Aken et al., 2010). However, phytoremediation can take longer to achieve remedial targets than conventional cleanup methods (Palmroth et al., 2002). Phytoremediation relies on plant root systems to foster the growth and activity of microorganisms through release of metabolic products, improved aeration and/or other mechanisms (Singer et al., 2003a). Increased microbial activity in plant-associated soils can lead to biodegradation of pollutants through either induction of microbial degradation pathways or co-metabolism. Some of the aromatic compounds released by plant roots share chemical similarities to pollutants, such as PHCs, or an intermediary in the metabolism of pollutants, such as salicylic acid. Plant aromatic compounds have been shown to induce the microbial degradation of petroleum (McCutcheon and Schnoor, 2003; Siciliano et al., 2003; Singer et al., 2003b). Co-metabolism of pollutants by soil microbial communities can occur when secondary compounds that are structurally analogous to contaminants are available to the community (Donnelly et al., 1994; Gilbert and Crowley, 1997; Miya and Firestone, 2001). At high latitudes, such as in Alaska, many plant species have unusually high concentrations of secondary compounds, which are thought to have evolved as a defense against damage, such as from UV radiation or herbivory, and are therefore particularly attractive for investigations into microbial degradation of aromatic compounds (Bryant et al., 1991, 1994; Hadacek, 2002; Stark et al., 2008). With increased concentrations of secondary compounds at higher latitudes, increased rates of contaminant disappearance may be achievable, which would be advantageous for promoting contaminant biodegradation in the rhizosphere. Another important advantage of using plants native to higher latitudes is their adaptation to the extreme environmental conditions, such as low temperature and nutrient limitation.

In this study we investigated the long-term effects of phytoremediation on contaminant disappearance at a site in Fairbanks, Alaska, where a phytoremediation

experiment on soils contaminated with crude oil or diesel fuel was previously conducted (Reynolds and Koenen, 1997; Reynolds et al., 1997b; Reynolds et al., 1999). Experimental field plots were established and subjected to different fertilization and planting regimes, while petroleum disappearance was monitored for two years. The site then underwent colonization by local plants and natural attenuation for 15 years. In this study, we re-examine this field site 15 years later to assess the long-term effects of phytoremediation on colonization by native plants, their rhizosphere microbial communities and petroleum removal from soil. The data provide a novel, long-term assessment of the feasibility of rhizoremediation with native plants and soil microbial communities for achieving regulatory cleanup guidelines in cold regions.

BACKGROUND AND SITE DESCRIPTION

A phytoremediation field study was initiated in 1995 and 1996 at the Farmers Loop Permafrost Research Facility field site of the Army Corps of Engineers Cold Regions Research and Engineering Laboratory (ACE CRREL) (Reynolds and Koenen, 1997; Reynolds et al., 1997a; Reynolds et al., 1997b; Reynolds et al., 1999). The study sought to compare the effects of nutrients and vegetation on rhizosphere-enhanced bioremediation of PHC contaminated soils. Crude oil and diesel contaminated soils were used for the experiment: crude oil contaminated soils were collected from a gravel pad at a pump station on the Trans-Alaska pipeline, and diesel contaminated soils were collected during the removal of an underground storage tank. Soils were transported to the Farmers Loop Facility and placed, separately, in adjacent lined and bermed areas approximately 21 x 3 m, and 60 cm deep (Figure 4-1 and 4-2). Each of the soil piles was subdivided by wooden beams into seven treatment plots. For each soil the treatments included: three levels of rhizosphere enhancement and two nutrient levels. The three levels of rhizosphere enhancement included

unplanted (-P), annual ryegrass (*Lolium multiflorum,* 1P), or a mixture of annual ryegrass and arctared fescue (*Festuca rubra*, 2P). Grasses were chosen for their cold hardiness, rapid growth, and tolerance to low-fertility soils. The two nutrient levels were either control (no fertilizer, -F) or added nutrients (+F). Commercially available fertilizer (granular 20-20-10) was surface applied at approximately 620 g/m^2 of N, P, and K. The annual ryegrass and arctared fescue mixture was 1:1 by weight, and seeding was done by hand at approximately 10.8 g/m^2.

Triplicate soil samples were taken from each of the plots to sample for TPH at day 0 and after 54 days and 238 days of treatment (described below). After 238 days of treatment, TPH concentrations in both the plant + fertilizer treatment and soils without any amendments had decreased relative to the initial TPH concentrations (measured at day 0). The planted and fertilized treatment had significantly lower TPH concentrations than the untreated soils (Figure 4-3).

After completion of the initial phytoremediation project in 1996, the site was not actively managed for approximately 15 years. No nutrient or seed amendments were made during this time. In 2010-2011 the site was re-examined and our follow up study was conducted.

MATERIALS AND METHODS

Site Preparation and Sampling. A follow up study of the initial phytoremediation experiment conducted at the Farmers Loop Permafrost Research Facility field site was conducted in June 2011. The original soils and treatment plots remained intact and were used with minor modification to the sampling design. Each of the seven treatment plots per contaminant was subdivided into six 1 x 1.5 m^2 sections using twine. Two soil samples were collected from each of the sub-sections into both a glass jar with Teflon lined lids for TPH

analysis or two separate sterile plastic zip-top bags for microbial analyses. Soil samples were collected from four random points within each sub-section at a depth of 10 cm and homogenized in the sampling container. Samples were sieved through a sterile 2.5 mm sieve and stored at -80°C until analysis for molecular and TPH analysis or at 4°C for culture-based microbial assays.

Vegetation Characterization. All vegetation within each sub-section was counted and identified to the species level (Johnson et al., 1995). Percent cover of dominant vegetation, bare ground, and mosses were visually assessed. Tree height, basal diameter, and canopy cover were also measured for each tree at the site. Trees were defined as such when main-stem height was greater than 20 cm, and trees shorter than 20 cm were counted as "seedlings".

Analysis of TPH in Soils. In the initial study (1995), TPH was extracted by sonication with methylene chloride (EPA Method 3550B). Anhydrous Na_2SO_4 was added to the soils during extraction as a drying agent and extracts were then analyzed by GC-FID.

For the follow up study, total petroleum hydrocarbons (TPH) were analyzed by Alpha Analytical (320 Forbes Boulevard, Mansfield, MA). Soil samples were extracted and analyzed using a gas chromatograph equipped with a flame ionization detector (GC/FID). The temperature program and associated experimental conditions were optimized to obtain maximum resolution in an 80 min chromatographic run representative of hydrocarbons in the n-octane (C8) to n-tetracontane (C40) range. Qualitative evaluation of the sample was conducted by reviewing the sample chromatogram in conjunction with a chromatogram of an alkane series generated with the same chromatographic conditions (EPA Method 8015M). Quantitative determination of the sample hydrocarbon concentration was performed in accordance with EPA Method 8015M.

Analysis of Nutrients in Soils. Soil physical and chemical analyses were conducted at the University of Alaska Fairbanks, Palmer Center for Sustainable Living. Properties assessed were: pH, % loss on ignition (C), CEC, NO_3^-, NH_4^+, P, K, and soil particle size (Bremner, 1982; Day, 1965; Holmgren et al., 1977; Jackson, 1958; Michaelson et al., 1987; Peech, 1982).

Microbial Community Analyses. Total soil DNA was extracted using the FastDNA SPIN kit for soil (MP Biomedicals, Ohio, USA) following the manufacturer's instructions. DNA was eluted into 50 µL of water and stored at -20^{o}C until analysis. DNA concentrations were evaluated by measuring absorbance at 260 and 280 nm using a NanoDrop ND-1000 Spectrophotometer (Thermo Fischer Scientific, USA).

Bacterial community profiling using T-RFLP was performed on extracted total community DNA. PCR reactions were carried out using a 5' 6-carboxyfluorescien-labelled 27F and unlabeled 1392R primers with a total reaction volume of 25 µl. The PCR reaction contained final concentrations of 1 x PCR buffer, 3 mM $MgCl_2$, 200 µM dNTPs and 0.1 µM BSA (New England Biolabs, Beverly, MA, USA), 200 nM each primer, 0.5 U Taq DNA polymerase and 1 µl template DNA. Thermal cycler conditions were as follows: denaturation at 94^{o}C for 5 min; 30 cycles of 94^{o}C for 1 min, 58^{o}C for 1 min, and 72^{o}C for 1 min 40 s; and a final extension step of 72^{o}C for 10 min. PCR products were purified using the Qiaquick PCR purification kit (Qiagen, Valencia, CA, USA) and quantified using the Nanodrop ND-1000 spectrophotometer. Then 40 ng of the PCR product was digested for 3 h at 37°C using HhaI restriction enzyme, with *Pseudomonas stutzeri* DNA run as a control to verify complete digestion. After digestion, fragments were precipitated overnight using 1.5 µl of sodium acetate (0.75M), 6 µg of glycogen per reaction (molecular grade), and 47 µl ethanol and centrifuged at 14 000 r.p.m. for 35 min. Pellets were resuspended in 1.0 µl H2O, 0.5 µl of ROX-labelled MapMarker 1000 ladder (BioVentures, Murfreesboro, TN, USA) and 9.5 µl

deionized formamide, and then run on an ABI 3100 Genetic Analyzer. T-RFLPs were analyzed using GeneMapper software version 3.7 (Applied Biosystems, Foster City, CA, USA). Terminal restriction fragments less than 50 bp and above 1000 bp in size were excluded from analyses.

Statistical Analyses. Data were tested for normality by testing the residuals of each data set, and were found to be non-normally distributed. Vegetation and TPH data were analyzed individually to detect trends using appropriate non-parametric tests (NMDS, Kruskal-Wallis test). Relationships between vegetation and TPH were analyzed using regression analyses. Vegetation communities were tested using ANOVA. Microbial community data was analyzed using detrended correspondence analysis (DCA). All statistical analyses were conducted in PAST (Version 2.15, April 2012, Hammer and Harper 2001), significance was accepted at p=0.05 (95% confidence level) for all statistical analyses.

RESULTS AND DISCUSSION

Soil TPH concentrations were significantly lower in 2011 than in samples quantified in 1996, when contaminant levels were last quantified and published for the site (Figure 4-3). In both the diesel and crude oil contaminated soils in 2011, contaminant levels in all treatment plots had dropped by 80-95% of levels last measured in 1996, well below the ADEC cleanup level for diesel range organics (1000 mg/kg) (ADEC, 2012). This suggests that natural attenuation (no amendments), land farming or phytoremediation are all capable of achieving cleanup limits for hydrocarbon contaminated soils with initial concentrations of greater than 8300 mg/kg within ~15 years. The concentration of residual hydrocarbons remained higher in the crude oil contaminated soils than in the diesel contaminated soils, which could be because recalcitrant compounds constitute a greater portion of the crude oil than diesel fuels (Reynolds et al., 1997a), as well as the fact that the crude oil contaminated soil was more

nutrient poor, more coarse, and more acidic. For most PHCs, reduction rates are often initially rapid, followed by a period of much slower losses, similar to the observed effects here. These data demonstrate that, despite slow disappearance rates, TPH disappearance continues to occur and, for soils contaminated with up to 3422 mg/kg diesel, can reach clean up levels in less than two decades and that for near-surface contaminants, phytoremediation, land farming or natural attenuation may provide an effective low-cost treatment.

Results from vegetation surveys showed that the plant community had changed substantially from when the site was first established in 1995 (Figure 4-1 and 4-2). None of the originally planted grasses were found on the site, instead the plots have been colonized by both native and non-native Alaskan plants (Figure 4-4). The diesel-contaminated soils were more heavily colonized by plants than crude oil contaminated soils ($p=0.06$), and there were more and larger woody plants present in the diesel contaminated soil (Figure 4-4). This is likely due to the difference in soil type between the diesel and crude-oil contaminated soils. The two contaminated soils differed in soil texture: crude oil contaminated soils were gravel with a large grain size, while diesel contaminated soils were finer textured with more organic matter (Table 4-1). Coarser soils tend to have a lower cation exchange capacity (CEC) and less water and nutrient retention, which can mean more inhospitable conditions for plant and microbial growth (Bergamaschi et al., 1997; Fredlund et al., 1997; Lund et al., 1999). Fine-textured soils with higher soil organic matter can be a more favorable environment for plant colonization and microbial growth, conditions which could lead to increased contaminant degradation (Olk and Gregorich, 2006). Interestingly, the crude oil contaminated soils were more heavily colonized by non-native plants from the *Asteraceae* family, such as *Taraxacum officinale*, and from the *Fabaceae* family, such as *Vicia cracca* and *Trifolium hybridum*.

Current vegetation status also appears to be affected by the original treatment type: plots that were originally planted with grasses grouped together in ordination space (NMDS, Crude

Oil:R^2 = 0.7005, Stress = 0.1927; Diesel: R^2 = 0.6859, Stress = 0.1622, Figure 4-5). NMDS also revealed that plots with no original treatment are associated with increased numbers of native plants (e.g. *Salix sp.*,O*xytropis deflexa)* and that initially planted and fertilized plots were associated with increased numbers of non-native plants (e.g. *Agrimonia striatica, Taraxacum officinale*). Nutrient status and other soil parameters have different effects on vegetation depending on the contamination, and therefore soil, type (Table 4-1, Figure 4-5). Both nutrient levels and contaminant type have been found to affect vegetation survival and colonization (Chapin et al., 1986; Rohr et al., 2006). Fertilization may have had an effect on initial colonization and survivability of vegetation; however, this direct effect would only last until the fertilizer had either been depleted or dispersed through the site, typically less than a year. Although direct effects of the fertilizer would be fairly short-lived, it is possible that the fertilizer acted to prime plant succession or to allow for some seeds to colonize and establish in a particular treatment. Although the fertilizer may not remain in the soil over the long-term, turnover of nitrogen-rich plant biomass might also provide a long-term enhancement of available nitrogen (Ruess et al., 1996).

A critical question from the standpoint of bioremediation is how the initial vegetation and fertilization treatments affected contaminant degradation rates. It is important to note that one characteristic of petroleum biodegradation is the sequential loss or disappearance of petroleum components, in which more labile fractions are lost most rapidly while recalcitrant fractions remain and slowly degrade in weathered petroleum. As a result, biodegradation follows a somewhat asymptotic pattern rather than a linear progression. The initial 3-year study found increased rates of contaminant disappearance in plots that were both planted and treated with fertilizer within the first year of the field study (Reynolds and Koenen, 1997; Reynolds et al., 1997b; Reynolds, 2004). These rates were higher than in treatments that were only either planted or fertilized. Fifteen years later, significant differences in mean soil TPH

concentrations between planting/fertilization treatments were not detected when mean TPH levels were compared at the plot level. Addition of vegetation and fertilizer appeared to accelerate TPH biodegradation relative to unplanted, unfertilized treatments initially, and over the long term all treatments arrived at a similar plateau.

Further statistical evaluation of the relationships between detailed vegetation data and TPH revealed that increased woody vegetation counts affected TPH disappearance over the long term. There was a negative relationship between number of woody plants present and soil TPH concentrations across all plots (Figure 4-6: crude oil contaminated soils; $R^2 = 0.315$, $p<0.01$: Diesel contaminated soils; $R^2 = 0.452$, $p<0.01$). In diesel contaminated soils, the current number of trees is significantly different between the original treatment types ($p<0.001$) with significantly fewer trees in cells which were originally planted and fertilized ($p<0.05$) compared to untreated and unplanted cells. Crude oil contaminated soils did not have differences between number of trees on each plot ($p=0.06$). The data suggest that combined planting and fertilization accelerated initial diesel biodegradation rates, yet in the long term may slow remediation by limiting the establishment of trees in diesel-contaminated soil.

The dominant woody plant species present included willow (*Salix* spp.), Alaskan birch (*Betula neoalaskana*), white spruce (*Picea glauca*), and balsam poplar (*Populus balsamifera*) (Figure 4-4). Some northern tree populations (*Betula*) have been shown to have increased concentrations of phenolic secondary compounds with increasing latitude (Stark et al., 2008). Based on the secondary compound hypothesis (Robertson et al., 2010; Slater et al., 2011), this increasing secondary compound concentration could have resulted in increased degradation potential in the microbial community and a resulting decrease in soil TPH. Trees have more extensive root systems than some forbs and grasses, which may also contribute to rhizoremediation efficiency. *Salix* and *Populus* species are commonly used in remediation

studies and have been found to lower TPH and reduce toxicity in TPH contaminated soils (Newman and Reynolds, 2004; Palmroth et al., 2002). Willows (*Salix* spp.) are very common throughout the Alaskan interior and Northern latitudes. Willows were found in each of the plots at the Farmers Loop Site, with willows being found up to 4 m tall in the cell with the lowest soil TPH concentration. In lab and greenhouse studies, a willow native to interior Alaska (*Salix alaxensis*) was found to accelerate microbial biodegradation of diesel range organics (McFarlin, 2010) and PCBs (Slater et al., 2011) in spiked soils.

Because microbial biodegradation is an important mechanism for TPH removal from soils, we investigated the structure of the bacterial community in an effort to determine the effect of TPH and vegetation on microbial community dynamics. Terminal restriction fragment length polymorphism (T-RFLP) data indicate that the current microbial community structure varies depending on the original treatments implemented 15 years ago (Figure 4-7). The original treatment type resulted in different microbial community structures, much as it did for vegetation dynamics (Figures 4-1, 4-2 and 4-4). This suggests that the original treatment influenced vegetation structure and succession, which affected associated microbial communities by mechanisms such as rhizodeposition and the quantity and quality of aboveground and belowground litter. It has been shown that as trees age and the plant community matures, associated microbial populations change and may foster pollutant degrading microorganisms (Newman and Reynolds, 2004). It has also been found that with increasing quantities of vegetation there are increasing quantities of microorganisms on sites contaminated with PHCs (Margesin et al., 2007; Reynolds et al., 1999). The microbial data do not appear to indicate a significant relationship between contaminant concentration and bacterial community structure, nor a relationship between current nutrient status and bacterial community structure ($p>0.05$); however more studies should be conducted to further understand the relationship between soil parameters and the microbial community. Future

studies should focus on the microorganisms involved in actual biodegradation of the contaminant; for example, the Most Probable Number analysis of petroleum degraders would be directly indicative of the degradation potential at the site and would be valuable for further site assessment and long-term monitoring. Determining the identity of petroleum-degrading microorganisms present would also be valuable, since detection of previously characterized petroleum degraders can reveal environmental preferences that may suggest effective ways to enhance biodegradation and to monitor biodegradative populations in the long term.

The results of this study indicate that trees and shrubs have a positive effect on TPH biodegradation by biostimulating microbial communities in the root zone. Yet, the use of grasses in the initial stages of this study were also clearly beneficial, as has been shown previously (Phillips et al., 2009). However, the grasses applied were annuals, requiring repeated seeding, which is labor-intensive and expensive over the long term. Instead of planting with exotic annual plants, another phytoremediation option would be the use of native perennial grasses or forbs, which are adapted to local conditions and have been shown to survive contamination and to induce degradation of petroleum hydrocarbons in previous studies located in Northern latitudes (Ferrera-Rodríguez et al., 2012; Robson et al., 2003). Initial cultivation of grasses appeared to inhibit natural recruitment of woody plants that were associated with increased long-term contaminant loss. Trees and shrubs can also be advantageous in that they offer increased rooting depth and long-term growth as compared to grasses and forbs. Our findings suggest that a combined approach of initial seeding with grasses and planting with trees or shrubs, such as with native willow cuttings, may lead to increased and sustained contaminant disappearance rates at a site. In addition to promoting biodegradation in surface soils, grasses would also help to provide a living cap on contaminated soils by stabilizing soils and minimizing contaminant escapement from the site via blowing dust. Trees or shrubs would act to increase long term biodegradation rates at

deeper soil layers, provide soil stabilization, and may also provide added hydraulic control, preventing leaching of contaminants. More information regarding the interactions between native graminoids and woody plants planted together would be valuable for developing optimal long-term rhizoremediation strategies.

CONCLUSIONS AND FUTURE DIRECTIONS

This study investigated the long-term effects of phytoremediation on soils contaminated by crude oil and diesel fuel in interior Alaska. In the initial 3-year period post treatment, for both crude oil and diesel contaminated soils, rhizosphere enhancement with non-native grasses and nutrient amendments resulted in increased rates of TPH reduction relative to natural attenuation, nutrient additions alone, or plants alone (Reynolds et al., 1997a). Fifteen years later, we found that increased TPH disappearance appeared to be associated with increased numbers of trees and shrubs such as willow (e.g. *Salix bebbiana, Salix alexensis, Salix glauca*), birch (*Betula neoalaskana*), white spruce (*Picea glauca*), and balsam poplar (*Populus balsamifera*), all native to the region, which had colonized the site. The original field treatments appeared to affect the total vegetation present at the site, the plant successional trajectory, and the soil bacterial community structure. All plots achieved soil TPH levels that were below regulatory cleanup limits within 15 years. Phytoremediation and fertilization accelerated biodegradation initially as shown in the three-year study using non-native annual grasses, and our long-term study revealed that colonization with trees and shrubs corresponded to increased extent of long-term petroleum removal. The original planting and fertilization strategy was effective for accelerating the reduction of contaminant levels and mitigating risk in a timely manner during the initial phase after contamination of a site. The increased TPH disappearance in plots with more trees and shrubs implies that initially planting with woody plants in addition to grasses may have the potential to achieve

TPH reductions greater than those observed here. Native vegetation may be the best phytoremediation option for ensuring long term survival of plants and the associated increase in degradation rates with minimal labor and expense.

Phytoremediation appears to be an effective and potentially low-cost bioremediation strategy particularly applicable to remote sites that are inaccessible to heavy equipment, that have no infrastructure, or that are situated on fragile ecosystems. The use of local plants like willows, which are readily propagated from stem cuttings, can further reduce costs associated with phytoremediation activities. Phytoremediation represents an option for the many contaminated sites located in Northern communities where infrastructure and resources limit the feasibility of conventional remediation strategies.

ACKNOWLEDGMENTS

The project described was supported by Grant Number 5P20RR016466 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH) and a graduate fellowship from the Alaska Idea Network for Biomedical Research Excellence (INBRE), and EPSCoR NSF award #EPS-0701898 & the State of Alaska. We thank Jan Fišer, Iva Pacovská from the Institute for Chemical Technology Prague for help with lab work. We also thank Chris Kasanke and Carl Richmond for help with field work and sample processing. Earlier site and laboratory work were supported by Army Environmental Quality Technology Program and the Alaska Science and Technology Foundation.

REFERENCES

ADEC, 2012. Oil and Other Hazardous Substances Pollution Control - 18 AAC 75, Psychosomatics. Juneau, AK.

Aken, B. Van, Correa, P.A., Schnoor, J.L., 2010. Phytoremediation of polychlorinated biphenyls: new trends and promises. Environ. Sci. Technol. 44, 2767–76.

Bergamaschi, B.A., Tsamakis, E., Keil, R.G., Eglinton, T.I., Montluçon, D.B., Hedges, J.I., 1997. The effect of grain size and surface area on organic matter, lignin and carbohydrate concentration, and molecular compositions in Peru Margin sediments. Geochim. Cosmochim. Acta 61, 1247–1260.

Bremner, J.M., 1982. Extraction of Exchangeable Ammonium, Nitrate, and Nitrite., in: Methods of Soil Anaysis, Part 2. Soil Science Society of America.

Bryant, J.P., Provenza, F.D., Pastor, J., Reichardt, P.B., Clausen, T.P., du Toit, J.T., Thomas, P., 1991. Interactions Between Woody Plants and Browsing Mammals Mediated by Secondary Metabolites. Annu. Rev. Ecol. Syst. 22, 431–446.

Bryant, J.T.P., Swihart, R.K., Reichardt, P.B., Newton, L., Biography, L., 1994. Biogeography of Woody Plant Chemical Defense against Snowshoe Hare Browsing: Comparison of Alaska and Eastern North America. Oikos 70, 385–395.

Chapin, F.S., Vitousek, P.M., Cleve, K. Van, 1986. The Nature of Nutrient Limitation in Plant Communities. Am. Nat. 127, 48–58.

Day, P.R., 1965. Particle Fractionation and Particle-Size Analysis, Hydrometer Method, in: Methods of Soil Anaysis, Part 1. Soil Science Society of America, Madison, WI, pp. 562–566.

Donnelly, P., Hegde, R.S., Fletcher, J.S., 1994. Growth of PCB-degrading bacteria on compounds from photosynthetic plants. Chemosphere 28, 981–988.

Ferrera-Rodríguez, O., Greer, C.W., Juck, D., Consaul, L.L., Martínez-Romero, E., Whyte, L.G., 2012. Hydrocarbon-degrading potential of microbial communities from Arctic plants. J. Appl. Microbiol. 1–13.

Filler, D., Reynolds, C., Snape, I., Daugulis, A.J., Barnes, D.L., Williams, P.J., 2006. Advances in engineered remediation for use in the Arctic and Antarctica. Polar Rec. (Gr. Brit). 42, 111–120.

Fredlund, M.D., Fredlund, D.G., Wilson, G.W., 1997. Prediction of the soil-water characteristic curve from grain-size distribution and volume-mass properties, in: 3rd Brazilian Symposium on Unsaturated Soils. pp. 1–12.

Gilbert, E.S., Crowley, D.E., 1997. Plant compounds that induce polychlorinated biphenyl biodegradation by Arthrobacter sp. strain B1B. Appl. Environ. Microbiol. 63, 1933–8.

Hadacek, F., 2002. Secondary Metabolites as Plant Traits: Current Assessment and Future Perspectives. CRC. Crit. Rev. Plant Sci. 21, 273–322.

Hammer, O., Harper, D.A.., 2001. PAST: Paleontological Statistics software package for education and data analysis. Palaeontol. Electron. 9.

Holmgren, G.S., Juve, R.L., Geschwender, R.C., 1977. A Mechanically Controlled Variable Rate Leaching Device. Soil Sci. Soc. Am. Journal1 41, 1207–1208.

Jackson, M.., 1958. Soil Chemical Analysis. Prentice Hall.

Johnson, D., Kershaw, L., MacKinnon, A., Pojar, J., 1995. Plants of the Western Forest: Alaska to Minnesota Boreal and Aspen Parkland. Lone Pine Publishing, Auburn.

Lund, E., Christy, C., Drummond, P., 1999. Practical applications of soil electrical conductivity mapping, in: Precision Agriculture. pp. 1–9.

Margesin, R., Hämmerle, M., Tscherko, D., 2007. Microbial activity and community composition during bioremediation of diesel-oil-contaminated soil: effects of hydrocarbon concentration, fertilizers, and incubation time. Microb. Ecol. 53, 259–69.

McCutcheon, S.C., Schnoor, 2003. Phytoremediation: Transformation and control of contaminants. Wiley & Sons Inc., New Jersey.

McFarlin, K., 2010. Rhizoremediation of diesel-contaminated soil using Salix alaxensis. University of Alaska Fairbanks.

Michaelson, G.J., Ping, C.L., Mitchell, G.A., 1987. Correlation of Mehlich 3, bray 1, and ammonium acetate extractable P, K, Ca, and Mg for Alaska agricultural soils 1. Commun. Soil Sci. Plant Anal. 18, 1003–1015.

Miya, R.K., Firestone, M.K., 2001. Enhanced phenanthrene biodegradation in soil by slender oat root exudates and root debris. J. Environ. Qual. 30, 1911–8.

Newman, L.A., Reynolds, C.M., 2004. Phytodegradation of organic compounds. Curr. Opin. Biotechnol. 15, 225–30.

Olk, D.C., Gregorich, E.G., 2006. Overview of the Symposium Proceedings, "Meaningful Pools in Determining Soil Carbon and Nitrogen Dynamics". Soil Sci. Soc. Am. J. 70, 967–974.

Palmroth, M.R.T., Pichtel, J., Puhakka, J.A., 2002. Phytoremediation of subarctic soil contaminated with diesel fuel. Bioresour. Technol. 84, 221–8.

Paudyn, K., Rutter, A., Kerry Rowe, R., Poland, J.S., 2008. Remediation of hydrocarbon contaminated soils in the Canadian Arctic by landfarming. Cold Reg. Sci. Technol. 53, 102–114.

Peech, M., 1982. Hydrogen Ion Activity, in: Methods of Soil Anaysis, Part 1. Soil Science Society of America, Madison, WI, pp. 920–921.

Phillips, L. a., Greer, C.W., Farrell, R.E., Germida, J.J., 2009. Field-scale assessment of weathered hydrocarbon degradation by mixed and single plant treatments. Appl. Soil Ecol. 42, 9–17.

Reynolds, C.M., 2004. Cold Regions Research Field Demonstration of Rhizosphere-Enhanced Treatment of Organics-Contaminated Soils on Native American Lands with Application to Northern FUD Sites.

Reynolds, C.M., Koenen, B.A., 1997. Rhizosphere-Enhanced Bioremediation. Mil. Eng. 586, 32–33.

Reynolds, C.M., Koenen, B.A., Carnahan, J.B., Walworth, J.L., Bhunia, P., 1997a. Rhizosphere and Nutrient Effects on Remediating Subarctic Soils, in: In Situ and On-Site Bioremediation: Vol 1. pp. 297–302.

Reynolds, C.M., Koenen, B.A., Perry, L.B., Pidgeon, C.S., 1997b. Initial Field Results for Rhizosphere Treatment of Contaminated Soils in Cold Regions, in: Zubeck, H.K., Woolard, C.R., White, D.M., Vinson, T.S. (Eds.), International Association of Cold Regions Development Studies. Anchorage, AK, USA, pp. 143–146.

Reynolds, C.M., Wolf, D.C., Gentry, T.J., Perry, L.B., Pidgeon, C.S., Koenen, B.A., Rogers, H.B., Beyrouty, C.A., 1999. Plant enhancement of indigenous soil micro-organisms : a low-cost treatment of contaminated soils. Polar Rec. (Gr. Brit). 35, 33–40.

Robertson, S.J., Kennedy, N.M., Massicotte, H.B., Rutherford, P.M., 2010. Enhanced biodegradation of petroleum hydrocarbons in the mycorrhizosphere of sub-boreal forest soils. Environ. Microbiol. Rep. 2, 587–593.

Robson, D.B., Knight, J.D., Farrell, R.E., Germida, J.J., 2003. Ability of cold-tolerant plants to grow in hydrocarbon-contaminated soil. Int. J. Phytoremediation 5, 105–123.

Rohr, J.R., Kerby, J.L., Sih, A., 2006. Community ecology as a framework for predicting contaminant effects. Trends Ecol. Evol. 21, 606–13.

Ruess, R.W., Cleve, K. Van, Yarie, J., Viereck, L.A., 1996. Contributions of fine root production and turnover to the carbon and nitrogen cycling in taiga forests of the Alaskan interior. Can. J. For. Res. 26, 1326–1336.

Sanscartier, D., Zeeb, B., Koch, I., Reimer, K., 2009. Bioremediation of diesel-contaminated soil by heated and humidified biopile system in cold climates. Cold Reg. Sci. Technol. 55, 167–173.

Siciliano, S.D., Germida, J.J., Banks, K., Greer, C.W., 2003. Changes in microbial community composition and function during a polyaromatic hydrocarbon phytoremediation field trial. Appl. Environ. Microbiol. 69, 483–9.

Singer, A., Crowley, D.E., Thompson, I.P., 2003. Secondary plant metabolites in phytoremediation and biotransformation. Trends Biotechnol. 21, 123–130.

Singer, A.C., Smith, D., Jury, W.A., Hathuc, K., Crowley, D.E., 2003. Impact of the plant rhizosphere and augmentation on remediation of polychlorinated biphenyl contaminated soil. Environ. Toxicol. Chem. 22, 1998–2004.

Slater, H., Gouin, T., Leigh, M.B., 2011. Assessing the potential for rhizoremediation of PCB contaminated soils in northern regions using native tree species. Chemosphere 84, 199–206.

Stark, S., Julkunen-Tiitto, R., Holappa, E., Mikkola, K., Nikula, A., 2008. Concentrations of foliar quercetin in natural populations of white birch (Betula pubescens) increase with latitude. J. Chem. Ecol. 34, 1382–91.

Uhlik, O., Jecna, K., Mackova, M., Vlcek, C., Hroudova, M., Demnerova, K., Paces, V., Macek, T., 2009. Biphenyl-metabolizing bacteria in the rhizosphere of horseradish and bulk soil contaminated by polychlorinated biphenyls as revealed by stable isotope probing. Appl. Environ. Microbiol. 75, 6471–7.

Wenzel, W.W., 2008. Rhizosphere processes and management in plant-assisted bioremediation (phytoremediation) of soils. Plant Soil 321, 385–408.

Yergeau, E., Lawrence, J.R., Sanschagrin, S., Waiser, M.J., Korber, D.R., Greer, C.W., 2012. Next-generation sequencing of microbial communities in the Athabasca River and its tributaries in relation to oil sands mining activities. Appl. Environ. Microbiol. 78, 7626–7637.

FIGURES

Figure 4-1: Photo overview of the study site in 1996 (top) and 2011 (bottom). Note tree establishment and colonization, and variety of vegetation colonizing the different plots.

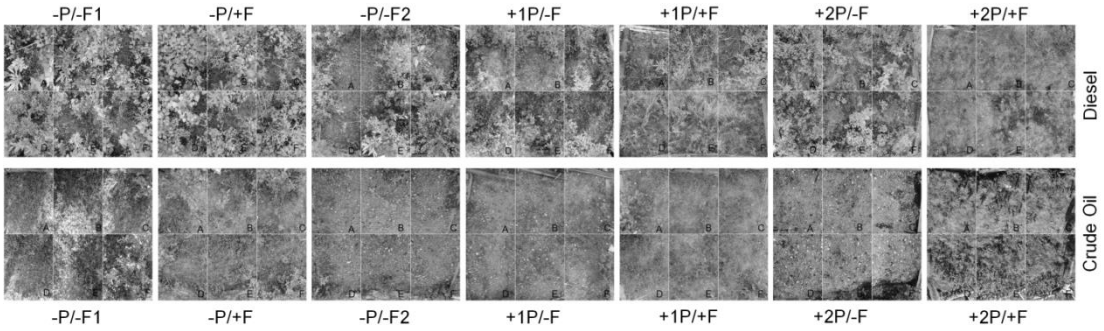

Figure 4-2: Aerial view of the design of the phytoremediation treatments at the site and the vegetation 15 years following establishment of the plots. Fourteen individual plots are present; half of the soils are contaminated with diesel fuel (top) and half with crude oil (bottom). Plots were planted with annual ryegrass (+1P) or a mix of annual ryegrass and arctared fescue (+2P), and treated with additional nutrients (+F1/+F2) or no added nutrients (-F).

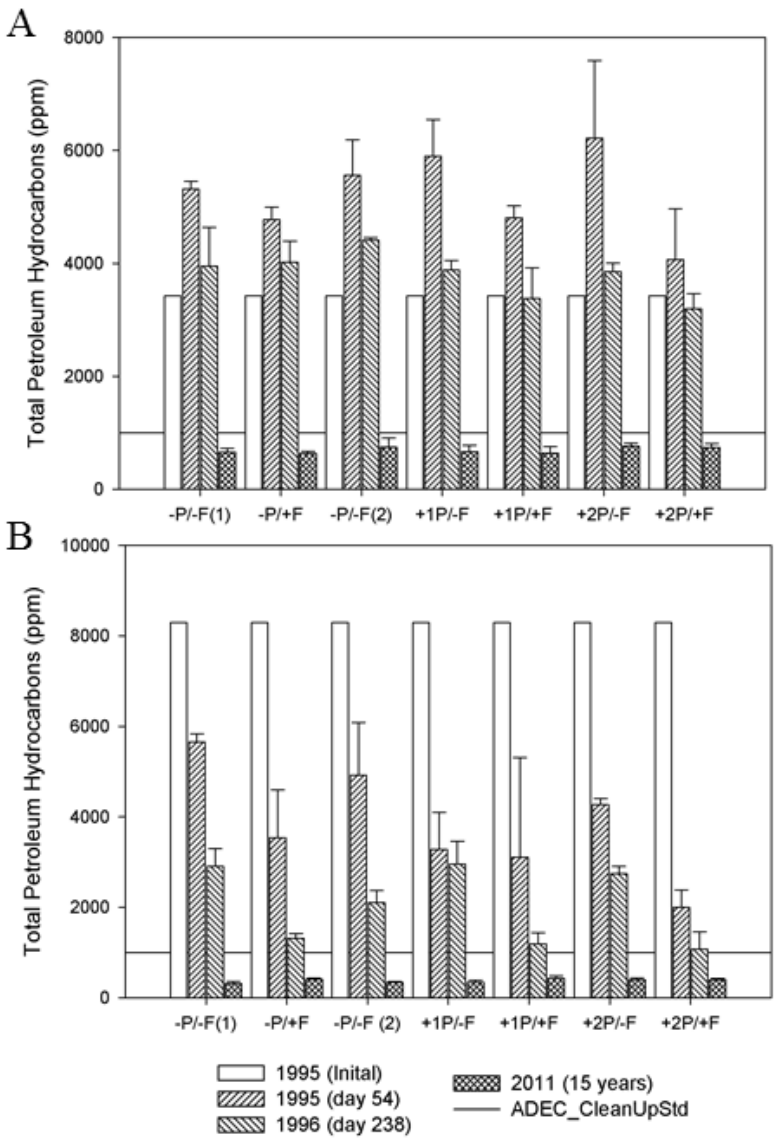

Figure 4-3: Changes in TPH concentrations in crude oil (A) and diesel (B) contaminated soils. Error bars indicated 95% confidence interval. Plots were planted with annual ryegrass (+1P) or a mix of annual ryegrass and arctared fescue (+2P), and treated with additional nutrients (+F) or no added nutrients (-F (1) or -F (2)). Alaska Department of Environmental Conservation clean-up standard for this site is indicated at 1000 ppm.

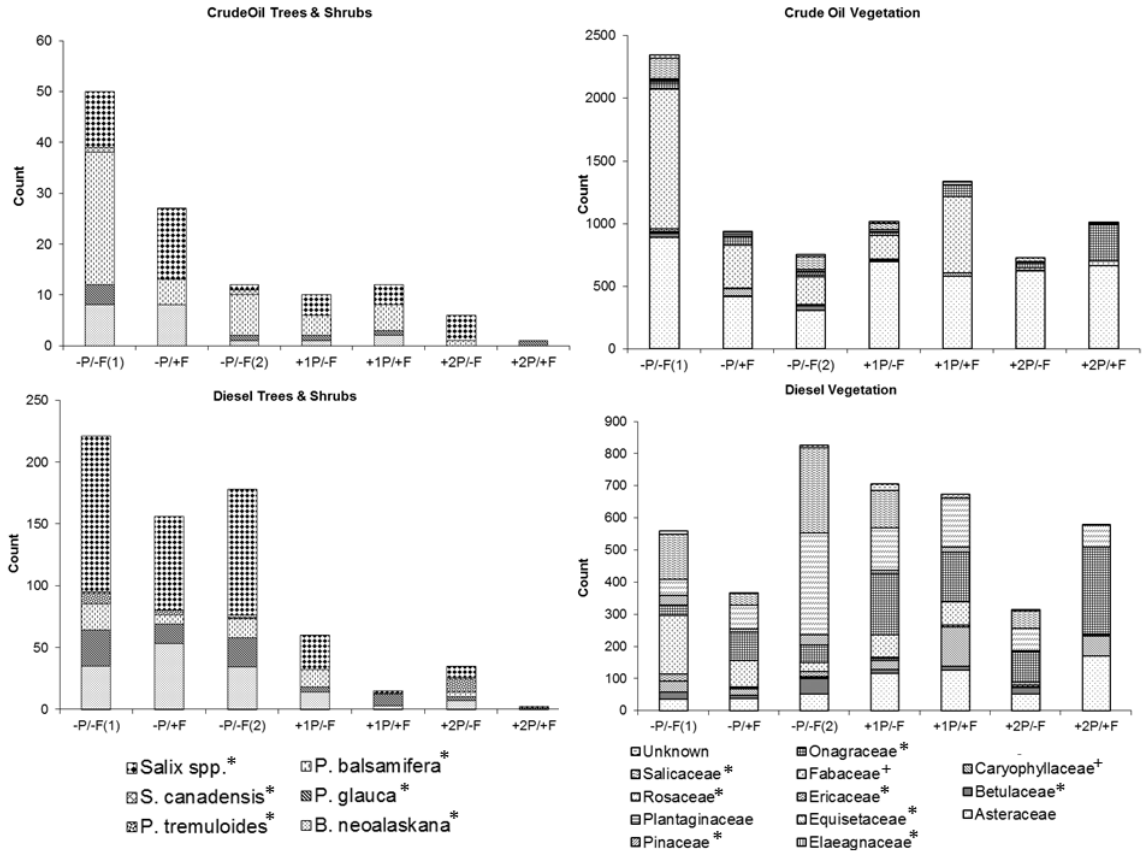

Figure 4-4: Total vegetation counts for crude oil (trees: top left, other vegetation: bottom left) and diesel (trees: top right, other vegetation: bottom right) contaminated soils. Plots were originally planted with annual ryegrass (+1P) or a mix of annual ryegrass and arctared fescue (+2P), and treated with additional nutrients (+F) or no added nutrients (-F (1) or -F (2)). Names designated with a (*) indicate native vegetation, names designated with a (+) indicate mixed native and non-native vegetation within a family, vegetation family names with no indicator are non-native.

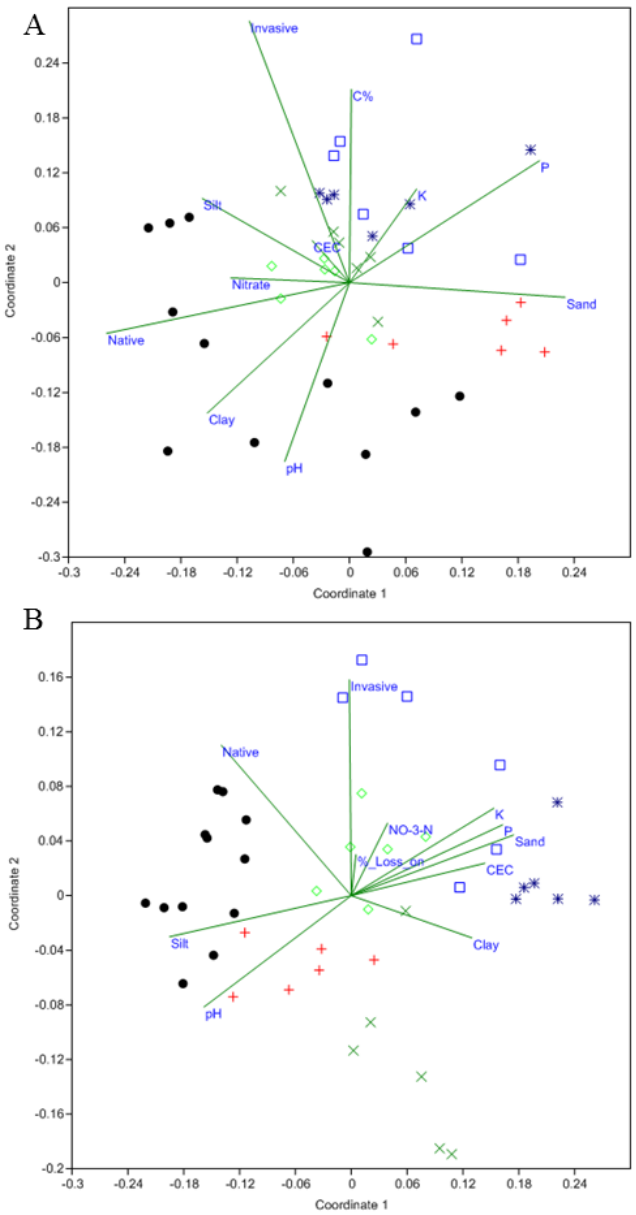

Figure 4-5: Nonmetric multidimensional scaling (NMDS) of vegetation data from crude oil (A) or diesel (B) contaminated soils. Blue points indicate plots planted and fertilized, green points are plots only planted, red points are plots only fertilized, and black points are plots with no added amendments. Vectors indicate environmental parameters.

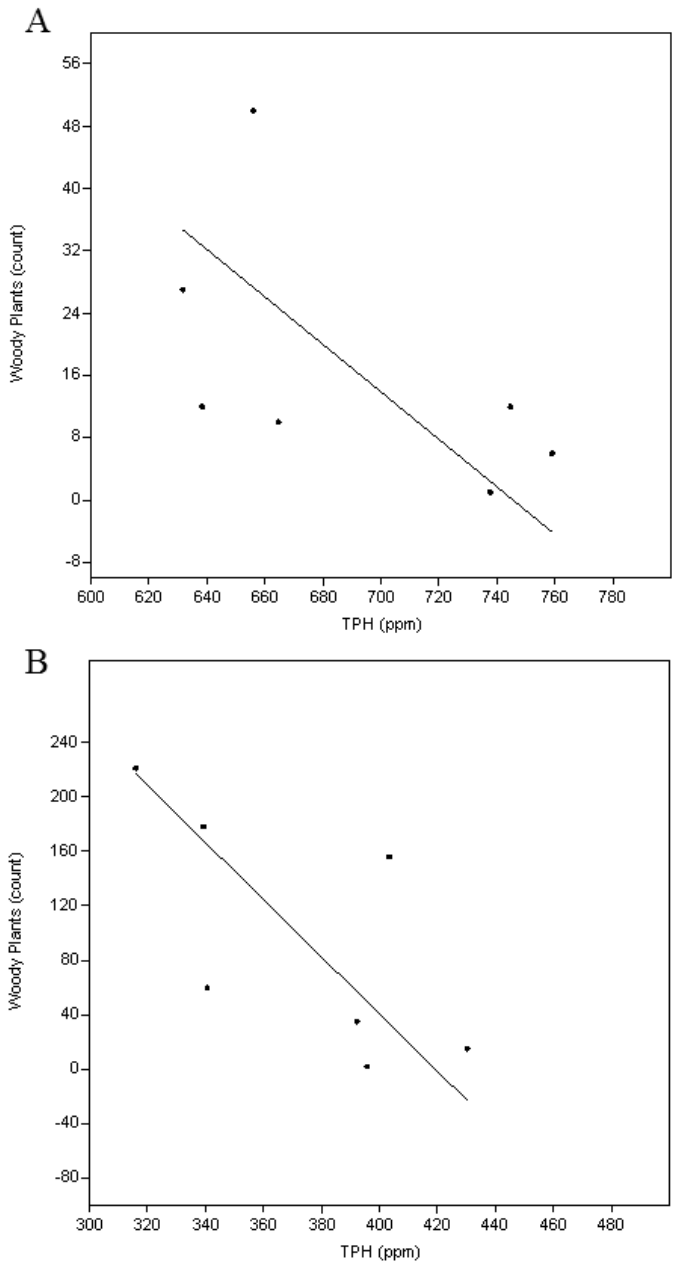

Figure 4-6: Regression of TPH concentrations and counts of all woody trees and shrubs for crude oil (A) or diesel (B) contaminated soils.

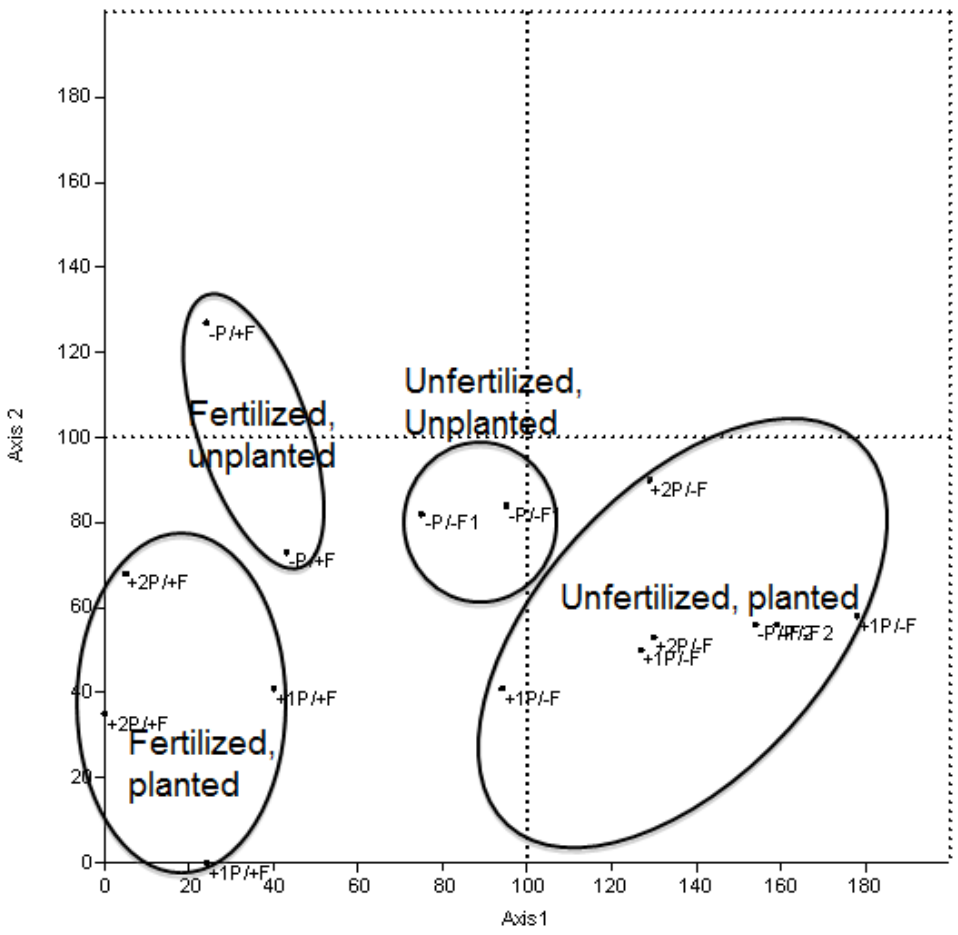

Figure 4-7: Detrended correspondence analysis (DCA) of T-RFLP of 16S rRNA for all samples from the crude oil contaminated soils. The first two axes account for 30% of total variance

TABLE

Cell Number	Original Treatment	Contaminant	pH	NH^{+4} (ppm)	NO_3^- (ppm)	P (ppm)	K (ppm)	CEC (meq/100)	C (%)	Sand (%)	Silt (%)	Clay (%)
1	-P/-F1	Crude Oil	6.9 ± 0.1	<1	30.1 ± 5.02	2.68 ± 0.58	50.85 ± 4.2	2.6 ± 0.2	1.3 ± 0.3	75.7 ± 5	15.8 ± 5	8.5 ± 0.6
2	-P/+F	Crude Oil	6.8 ± 0.2	<1	17.7 ± 6.52	134.1 ± 40.2	58.18 ± 8.9	2.4 ± 0.2	1.2 ± 0.1	80.5 ± 1	12 ± 0.5	7.5 ± 0.6
3	-P/-F2	Crude Oil	7.1 ± 0	<1	4.68 ± 2.52	1 ± 0	35.44 ± 4.06	2 ± 0.1	0.9 ± 0.1	78.9 ± 1.5	14.3 ± 1.2	6.8 ± 1
4	+1P/-F	Crude Oil	7.3 ± 0.1	1	2.67 ± 2.9	1.67 ± 0.58	32.43 ± 2.53	1.9 ± 0.3	0.9 ± 0.1	80.4 ± 2.1	13.5 ± 1.5	6.1 ± 0.6
5	+1P/+F	Crude Oil	6.6 ± 0.1	1	17.4 ± 8.11	89.99 ± 24.9	57.54 ± 2.53	2.5 ± 0.1	1.5 ± 0.1	79 ± 2.8	15 ± 2.5	6 ± 0.3
6	+2P/-F	Crude Oil	7.1 ± 0.1	<1	4.51 ± 0.71	1.34 ± 0.58	26.4 ± 2.08	1.7 ± 0.1	1 ± 0.1	77.6 ± 0	17.3 ± 0.6	5.1 ± 0.6
7	+2P/+F	Crude Oil	6.7 ± 0.1	1	3.51 ± 0.71	158.8 ± 51.4	55.5 ± 9.1	2.1 ± 0.1	1.3 ± 0	79.6 ± 2	14.6 ± 2	5.8 ± 0
8	-P/-F1	Diesel	6 ± 0.5	<1	32.5 ± 6.46	4.02 ± 1.74	60 ± 7.63	4.7 ± 0.4	2.8 ± 0.3	77.5 ± 1.9	19.3 ± 1.5	3.1 ± 0.6
9	-P/+F	Diesel	5.1 ± 0.1	<1	40 ± 8.45	91.31 ± 19.7	62.11 ± 10.8	5.6 ± 0.8	3.5 ± 0.4	78.1 ± 0.2	18.4 ± 0.7	3.5 ± 0.6
10	-P/-F2	Diesel	5.5 ± 0	<1	13.8 ± 2.1	2.01 ± 0	53.02 ± 9.37	4.9 ± 0.6	3.2 ± 0.5	80.6 ± 2.7	16.3 ± 2.7	3.1 ± 0.2
11	-P/-F2	Diesel	5.4 ± 0.1	<1	18.5 ± 6.55	2.01 ± 0	46.66 ± 4.98	5 ± 0.3	3.5 ± 0.2	82 ± 0.8	15.4 ± 1.4	2.6 ± 0.7
12	+1P/+F	Diesel	4.7 ± 0	<1	80.9 ± 27.9	95.64 ± 6.34	79.53 ± 16.5	5.3 ± 0.5	3.2 ± 0.5	81.1 ± 1.2	14.5 ± 1	4.5 ± 0.2
13	+2P/-F	Diesel	5.4 ± 0.1	<1	40.6 ± 46.2	2.68 ± 1.16	51.68 ± 12.8	4.9 ± 0.7	3.1 ± 0.2	81.7 ± 0.5	14.1 ± 0.8	4.3 ± 0.6
14	+2P/+F	Diesel	5.1 ± 0.2	<1	17.8 ± 11.6	161.1 ± 52	115.1 ± 28	6.1 ± 0.7	3 ± 0.3	82.9 ± 1.5	13.1 ± 1.5	3.9 ± 0.6

Table 4-1: Physical and chemical characteristics of soils from each plot.

Chapter 5

Conclusions

This research investigated the potential mechanisms and effectiveness of microbial communities and native boreal vegetation associated with contaminant degradation (i.e. disturbance) and biogeochemical cycling. This research investigated the effects of disturbance (e.g. contamination) and potential disturbance (genetic potential for contaminant degradation) on microbial interactions with the environment. This study shows that , in a PCB contaminated environment, the microbial community acts in concert to degrade PCBs and its degradation products. In addition, it was found that dominant vegetation type in a non-contaminated environment is the dominant factor that drives microbial community structure and functional potential, and that certain plant communities have the capacity to biostimulate microbial degradation of xenobiotic contaminants. Finally, I demonstrated that phytoremediation and fertilization treatments have a long term effect on contaminant disappearance, microbial communities, and vegetation in boreal forest soil, and that treatment strategies involving native perennial grasses and trees may be a practical and successful strategy for phytoremediation of petroleum-contaminated soils. Together, the results provide novel insights into the mechanisms of contaminant removal in boreal forest soils, the role of plants in ecosystem resilience to contamination, and demonstrate that phytoremediation using native and local plants can be an effective means to treat petroleum-contaminated soils.

In order to capitalize upon the functional capacity of microorganisms to remove organic chemicals from the environment, it is important to understand how the microbial community interacts with environmental inputs and disturbance. Here I have conducted investigations into the biodegradation potential of indigenous microbial communities, and the

mechanisms by which plants may biostimulate contaminant biodegradation by microorganisms. Yet, further investigations into the mechanisms that drive this functional potential are needed. Plant secondary metabolites (PSM) may play a critical role in microbially-mediated ecosystem processes. Understanding the fundamental chemical ecology of microbe-PSM interactions and decomposition of complex aromatics (e.g. organic matter and contaminants) is valuable to advancing phytoremediation technologies and is critical to understanding environmental controls on biogeochemical cycling, since PSMs have been hypothesized to promote contaminant biodegradation but inhibit organic matter decomposition (1, 2). Future studies should include phytochemical analyses and microbial functional assays, such as metatranscriptomics and studies of biodegradation and decomposition to quantify microbial function associated with boreal vegetation in contaminated and non-contaminated sites. A complete understanding of the chemical ecology of boreal forest plant-microbe interactions could pave the way for successful phytoremediation technologies and improved models for predicting ecosystem resilience and response to disturbance.

REFERENCES

1. **Singer AC, Smith D, Jury WA, Hathuc K, Crowley DE**. 2003. Impact of the plant rhizosphere and augmentation on remediation of polychlorinated biphenyl contaminated soil. Environ. Toxicol. Chem. **22**:1998–2004.

2. **Singer A, Crowley DE, Thompson IP**. 2003. Secondary plant metabolites in phytoremediation and biotransformation. Trends Biotechnol. **21**:123–130.

Appendix 1

Approaches to Microbial Diversity Analysis in Contaminated Environments[1]

ABSTRACT

Recent findings that only a small fraction of microbes have yet to be cultivated have resulted in the development of many novel approaches for studying the overall microbial species and metabolic diversity. These approaches are aimed at revealing and exploiting the genetic and metabolic potential that has remained hidden. Nowadays, scientists are capable of not only resolving the genetic code and taxonomically classifying individual microorganisms, but also ascribing the function of microorganisms in relation to their genetic material and, in turn, identifying those microbial populations that play key roles in habitat functioning. These advances have been enabled by the improvement of molecular biological methods, such as DNA isolation directly from the environment, polymerase chain reaction (PCR), DNA cloning and sequencing, community fingerprinting techniques, or stable isotope probing, as well as analytical methods. A huge recent development of bioinformatics and the existence of public databases, including Ribosomal Database Project (RDP) or Functional Gene Pipeline/Repository, and available applications, e.g. Molecular Evolutionary Genetics Analysis (MEGA), MOTHUR, FastGroupII, and many others, have covered the methodological advances by allowing for appropriate interpretation of obtained results. Microbial ecology can benefit from such progress by understanding the function of microbes not only in the context of their importance for xenobiotic transformation but also their function in terms of nature as a whole. The aim of this chapter is to discuss the approaches

[1] Uhlik, O., Leewis, M.C., Kurzawova, V., Lovecka, P., Stursa, P., Demnerova, K., Mackova, M., Macek, T. "Chapter 4: Approaches to Microbial Diversity Analysis in Contaminated Environments" in Trends in Bioremediation and Phytoremediation. Chemosphere. (Edited by Dr. G Plaza). 2010. ISBN: 978-81-308-0424-8.

Note: Leewis and Uhlik contributed equally to this work.

that can be used for the analyses of microbial, especially bacterial, diversity in contaminated environments.

INTRODUCTION

Microorganisms, especially bacteria, represent the beginning as well as the end of the food chain, and microbial processes act as a foundation in global nutrient cycling. Therefore, all organisms in the biosphere are actually dependent on microbial activity. The state of equilibrium of microbial diversity, however, is often affected by anthropogenic activities such as environmental pollution. In order to address how changes in microbial community structure in response to environmental pollution influence aboveground ecosystems and their functions, reliable ways of studying soil microorganisms and their diversity in particular environments are necessary.

Bacteria are an important part of environmental biota not only because of their species diversity (species richness and species evenness) but also because of the variety of metabolic activities (i.e. metabolic diversity) which they perform [1]. The extraordinary diversity and habitat range displayed by bacteria has been likely to be cause by some 3.8 billion years of evolution, which is 2 billion years longer than eukaryotic organisms [2]. The greatest bacterial diversity sources are considered to be pristine soils and sediments with high organic content, which have been estimated to contain over 11,000 different genomes per gram, i.e. about 10^4 different prokaryotic species, none of which is dominant [2,3]. But surprisingly, although such habitats have high species diversity, they have below-average phylogenetic diversity, which is likely to be caused by the presence of closely related species [4]. In contrast, environmental perturbations, such as pollution, result in a strong reduction of diversity and to selective advantage to certain populations which can adapt faster to changing conditions and therefore become dominant [3].

Whereas diversity analysis of microbial eukaryotes is still in its infancy, in the last few decades the widespread use of molecular biological methods has permitted researchers to inspect prokaryotic diversity, which in most environments is orders of magnitude larger than culture-based techniques are able to reveal [5,6]. The introduction of DNA-based molecular methods of diversity analysis has broadened researchers' view of microbial diversity, especially given that less than 1% of microbes have been cultivated under laboratory conditions [5,7,8]. The importance of accurate approaches to diversity determination for remediation research [9,10] is undisputable. Therefore, we will explore methods of analyzing bacterial diversity in the subsequent text

16S RIBOSOMAL RNA GENE

With the introduction of molecular-based methods in mid-1980s, the 16S rRNA gene (16S for short, rRNA for ribosomal RNA) has become the most important gene for prokaryote diversity studies [11]. Due to its uncommon properties namely ubiquity, extreme sequence conservation, domain structure with variable evolutionary rates, and not being a subject of horizontal transfer, its sequence knowledge permits identification to a certain level of both cultured and uncultured bacteria as well as the prediction of phylogenetic relationships [12,13]. In 1987 twelve major divisions (phyla) in the domain *Bacteria* were identified according to rRNA sequence data of almost all major cultured groups of bacteria that were accumulated during the previous century research. However, later culture-independent surveys revealed that there are at least 40 well-resolved major bacterial divisions, of which about 30 major divisions have no or very few cultured representatives [14].

BACTERIAL SPECIES DESIGNATION

With bacteria, as well as other microbes, binominal Linnaean names have been used to designate species, however, microbiologists lack a widely accepted theoretical species concept, and current methods for defining prokaryotic species do not adequately reflect the diversity that has been uncovered in nature. Instead, bacterial species are currently defined by an approach which is based on distinct rules for genotypic and phenotypic properties and is consistent with recent results from complete genome sequences [15,16]: (i) bacteria are assigned to a common species if their reciprocal pairwise DNA reassociation values are ≥70% in DNA-DNA hybridization experiments under standardized conditions, and their melting temperature difference is ≤5 °C; (ii) a species name is assigned only if its members can be distinguished from other species by at least one diagnostic phenotypic trait; (iii) all strains within a species must be to a certain degree phenotypically consistent, and (iv) species description should be based on more than one type strain. These rules for developing nomenclature, however, can be problematic in execution for a variety of reason. DNA-DNA hybridization is difficult to standardize between laboratories and it is inappropriate for rapid identification of prokaryotes. In addition, bacteria cannot be classified by DNA-DNA hybridization of a single strain by comparing it with a database using a set of criteria to assign the strain to a known taxon. Their classification requires comparisons of two genomes [16]. Depending on which DNA is used as target and which as probe, different values can be obtained with the same pair of strains [15]. Therefore, efforts are made to define species on the basis of average nucleotide identity (ANI) of all orthologous genes (i.e. genes that are similar to each other because they originated from a common ancestor) in complete genome sequences of strains. Data accumulated so far indicate that 70% DNA-DNA similarity correspond to 95% ANI and 69% conserved DNA [17].

In microbial ecology and environmental microbiology, the term species is typically substituted with operational taxonomic units (OTUs) – the number of distinct 16S rRNA sequences at a certain cut-off level of sequence diversity [15]. In practice, the cut-off threshold used is mostly 97% [16] because this value usually corresponds to 70% DNA-DNA similarity, i.e. the threshold for assigning bacteria to a common species. Yet, the correlation between 16S rRNA gene identity and DNA-DNA similarity values is highly variable in different lineages, and there are bacteria with >97% identity in the 16S rRNA genes that would be considered different species according to DNA-DNA similarity [8]. For these reasons, the term 'OTU' is usually preferred to the term 'species' for a cluster of related 16S rRNA sequences in environmental surveys.

BEYOND 16S RIBOSOMAL RNA GENE ANALYSES

Whereas bacteria that have less than 97% similarity of 16S rRNA gene sequences belong to different species, ≥97% sequence identity might not meet the 70% DNA-DNA hybridization criterion for inclusion in the same species. Therefore, in the majority of cases, 16S rRNA gene sequence or its portion assigns a bacterium to a genus but lacks resolution below that level [16]. The Ribosomal Database Project (RDP) has reported accuracy in the classification of partial 16S rRNA gene sequences to the genus level for 400-base reads and to the family level for 200-base reads [18]. A tool that might be useful in grouping bacteria into the major genetic lineages within a species is multilocus sequence typing (MLST) – the use of allelic mismatches of a small number (usually 7) of "housekeeping gene" loci, which evolve relatively slowly but more rapidly than 16S rRNA genes, and most of the variation that accumulates in these genes is considered selectively neutral [19]. Groups of isolates with identical profiles define strains or clones [16]. MLST was originally designed for molecular epidemiology and defines strains within already named species. In order to establish the

phylogenetic position of new species and the relationships between species in closely related genera, multilocus sequence analysis (MLSA) can been used [19]. This employs sequencing of multiple protein-coding genes (e.g. housekeeping genes) and subsequent comparison of the sequences [16].

ASSESSING BACTERIAL POPULATIONS AND THEIR DIVERSITY IN CONTAMINATED ENVIRONMENTS

Methods for analyzing bacterial populations and their diversity, with a focus on contaminated environments, are described below (Fig. A1-1).

(1) Culture-based methods

As outlined above, culture-based studies underestimate bacterial diversity due to two main reasons [14]. First, the current culturing methods target only those bacteria for which the proper conditions of cultivation have been found, whereas for the others, physical or biological factors have not been identified that would enable their laboratory cultivation; and second, researchers do not have the facilities to examine all the conditions for natural microbial samples, and therefore, the observed diversity on media may not be representative of those in nature. Although culture-based techniques present an incomplete picture of bacterial diversity, they are still widely used for their simplicity, cost-effectiveness, and mainly for the fact that cultivation still remains the best way to demonstrate metabolic capabilities of individual bacterial strains [20]. It should be noted however, there appears to be no reason why the formerly uncultured bacteria cannot be made culturable with different cultivation strategies [18].

Traditionally, after establishment of a pure culture, bacteria have been characterized and identified according to their phenotypic characters [21]. Phenotypic characteristics monitored include morphology, pigmentation, reaction to dyes, the presence or absence of

spores, nutrition requirements, sensitivity to inhibitors, and abilities of bacteria to assimilate, ferment, decarboxylate, or cleave selected organic compounds (i.e. biochemical activity). Determination of these characters as well as serological methods are reliable for bacterial identification in clinical samples, where a limited number of organisms occur. These methods are more difficult to sue with environmental samples because they are likely to contain many previously uncharacterized and uncultured strains, therefore the phenotypic characteristics as an identification tool are usually replaced by methods of molecular biology.

Recently, whole-cell MALDI-TOF methodology has been developed to identify bacteria with high reproducibility. This method identifies bacteria using a unique fingerprint of the most basic and most abundant proteins within the cell, typically the ribosomal proteins. In addition to being very abundant (about 20% of protein mass and about 3% of the total cellular mass), these proteins are also distinct between individual species or strains hence ideal for differentiation of cultured bacteria at the genus, species, and in some cases strain level [22]. Accurate and rapid identification of bacteria has been allowed due to the MALDI BioTyper developed by Bruker Daltonics. Applying this software enables one to compare the generated peak list with a library of characteristic spectra information of species and subspecies. The library spectra are generated by measurement of known bacterial strains at different days under slightly different conditions with subsequent generation of the list of typical peaks from the whole set of spectra. MALDI-TOF MS and a comprehensive database of well characterized species thus allows for very rapid and highly reproducible identification of bacteria. Finally, family trees generated by MALDI-TOF MS profile analyses are similar to trees based on 16S rRNA gene sequences [23,24].

(1.1) PLATE COUNTS

The oldest way of assessing bacterial diversity is extracting bacteria from a matrix, selectively plating the extracts and directly counting viable colonies which grow. Direct plating does not require any specialized equipment and is relatively simple to perform. Depending on the type of analysis, nutrient-rich complex, chemically defined, selective, differential, selective and differential, or minimal medium can be used – nutrient-rich media are aimed to assess diversity of the whole cultivable fraction, selective media allows only the growth of a few kinds of bacteria while inhibiting the growth of others, differential media do not inhibit any bacteria from growing, but the growth is biochemically or physically differentiated, and minimal media select for bacteria that are capable of metabolizing individual substrates, such as pollutants or their analogues, which are added as sole carbon and energy sources. In order to confirm that the bacteria truly utilized the target compound rather than impurities in agar or carryover from the soil suspension, secondary screening of colonies grown on minimal media is often desirable. For these purposes, several clearing-zone and colorimetric assays have been developed [20,25]. Cultivation of bacteria on agar plates also permits their further characterization from the point of view of taxonomical classification, or biochemical, physiological, and genetic properties. The plate count method is limited due to the fact that this method favors faster growing bacteria and, above all, that most bacterial species are unable to grow under commonly used incubation conditions [12]. In addition, bacterial cultivation is strongly dependent on the medium used and even nutrient-rich medium, which is designed to enable growth of most bacterial species, selects only for a small portion of bacterial types [3].

(1.2) LIQUID ENRICHMENT CULTURES

Enrichment methods are valuable for generating a mixed culture in which the portion of the population which uses a specific nutrient of interest, such as a pollutant, predominates. These cultures are typically constructed by inoculating mineral solution with a sample of soil, and a compound of interest is added to this solution as the sole carbon source. Incubation lasts until significant growth has occurred, generally when the medium becomes turbid. In order to achieve a stable consortium, several passages are often necessary. Liquid enrichment cultures are useful to produce isolates for further studying, preferably those with higher maximum specific growth rates; however, they do not reflect true diversity of culturable bacteria capable of degrading the compound of interest [20].

(1.3) SOLE CARBON SOURCE UTILIZATION PATTERNS

Sole carbon source utilization patterns allow for assessing potential metabolic diversity of bacterial communities [1]. In essence, a suspension of bacterial cells is diluted to extinction in a microtiter plate containing specific carbon substrates, including pollutants. Utilization of the substrate is detected by the reduction of an iodonitrotetrazolium dye (for substrates without pigmented products) or by the generation of colored products. The quantification of potential degraders is expressed as most probable number [20].

(2) Molecular biological methods based on DNA (RNA)

The use of molecular methods (based on studies of DNA isolated directly from the environment) for bacterial analyses was introduced in 1980s [26]. Since then and especially since the early 1990s when PCR-based analyses became widespread, much more accurate estimations of bacterial diversity compared to those obtained by culture-based techniques have been introduced [14].

(2.1) PCR-BASED METHODS

The general limitation of molecular methods is the fact that large quantities of DNA are required, which may be a problem in the case of isolation of total DNA from environmental samples. This problem can be overcome by amplifying target DNA by polymerase chain reaction (PCR) using specific primers [27]. However, PCR-based diversity studies can be biased by the formation of chimeras or heterologous duplexes, nonspecific priming of unexpected targets, the production of single stranded products, inhibition from compounds that copurify with nucleic acids [28], or different affinities of primers to templates and hence different amplification of target genes [12]. For instance, commonly used primers targeting 16S rRNA genes are known to miss some organisms owing to target mismatches. This limitation seems to be overcome by the use of 'miniprimers' – short (10-nucleotide) primers that expand the scope of detectable sequences without an increased rate of amplifying false positives [29]. Still, sequences with lower G+C content are thought to separate more efficiently in the denaturing step of PCR and can therefore be preferentially amplified [12].

Methods based on PCR include sequence analyses of clone libraries, gene quantification, and community fingerprinting techniques, namely terminal restriction fragment length polymorphism, denaturing or temperature gradient gel electrophoresis, temporal temperature gel electrophoresis, single-strand conformation polymorphism, and ribosomal intergenic spacer analysis. In general, the most commonly used target for community profiling has been the 16S rRNA gene, but also genetic polymorphism in other genes has been investigated [28].

Whole genome amplification (WGA) is a recently developed method for DNA multiplication which can overcome some of the bias issues encountered with traditional PCR. The WGA reaction is in principal similar to PCR, however it uses the Phi-29 polymerase to

amplify the entire genome [30]. This method can eliminate the requirement of adding synthetic primers and reliably amplifies the entire bacterial genome [31].

(2.1.1) Sequence analyses of clone libraries

Clone library construction of PCR amplified genes from DNA extracted directly from an environmental matrix followed by subsequent sequence analyses are powerful tools for investigating the diversity of bacterial communities. PCR targets either 16S rRNA or functional genes. Analyses of 16S rRNA gene libraries permit identification of the most dominant members of the communities as well as studying phylogenetic relatedness [1,11]. Analyses of functional gene libraries are connected with pollutant-degrading genes and their genetic diversity is connected with the substrate specificity [20].

(2.1.2) Quantitative PCR

Depending on the target genes, quantitative real-time PCR can provide a great deal of pertinent information in a rapid manner. 16S rRNA gene quantification is used for enumerations of bacteria and provides information on growth characteristics of bacterial communities, whereas catabolic (e.g. degradative) gene quantification gives one the insight into the metabolic capabilities of the community [20,32,33]. In real-time PCR, a PCR reaction is performed in a specially designed thermocycler in the presence of either non-specific fluorescent dye that binds to double-stranded DNA or alternatively fluorescent probes that hybridize to the specific target sequence. The fluorescence signal is measured at each amplification cycle, thus allowing the quantification relative to a standard curve containing known copy numbers of the target sequence [20].

(2.1.3) Terminal restriction fragment length polymorphism (T-RFLP)

T-RFLP is a molecular method for microbial community analysis in which an existing technology of restriction fragment length polymorphism (RFLP) has been further developed [34]. RFLP, also known as amplified rDNA restriction analysis (ARDRA) employs PCR to amplify a selected region of target gene from total community DNA and subsequent digestion of PCR products with restriction enzymes. The analysis of the restriction digest is typically performed on agarose gels. However, for community analysis the pattern becomes too complex and the large number of fragments cannot be resolved. T-RFLP is based on analysis of only terminal restriction fragments (T-RFs) and therefore is much more suitable for estimating the diversity of both phylogenetic and functional markers in complex community systems [35]. In T-RFLP, PCR amplification and restriction digest are similar to RFLP with the exception that one of the two primers used is fluorescently labeled at the 5'-end. The sizes of the terminal fluorescently labeled restriction fragments are determined by automated DNA sequencer with the limits of resolution for fragment size of 700 and 1000 bases with gel or capillary electrophoresis, respectively. If the target gene exceeds this size, two labeled primers with different fluorescent labels can be used for full coverage of the target. The type of attached fluorescent label is usually dictated by the detection technology of automated sequencers [28].

The most frequently used marker for T-RFLP analyses has been the 16S rRNA gene due to its huge phylogenetic utility and the existence of robust sequence databases. In addition, T-RFLP (or other community fingerprinting) analysis of 16S rRNA gene provides an approximate estimate of bacterial diversity. Although this method is PCR based and therefore subject to the previously discussed limitations, T-RFLP has the advantage that the sequence database can be used to assign detected populations to a possible taxon [28], and can be used to compare individual proportional representations.

(2.1.4) Denaturing and temperature gradient gel electrophoresis (DGGE, TGGE), temporal temperature gel electrophoresis (TTGE)

The original purpose of DGGE and TGGE was to detect point mutations in DNA sequences [36] but since then both techniques have been widely used for community profiling. The principal of these methods lies in decreased electrophoretic mobility of DNA in polyacrylamide gels containing a linear gradient of DNA denaturants in the case of DGGE or a linear temperature gradient in the case of TGGE. TTGE was first introduced as a modification of TGGE – in TTGE, the temperature of a gel plate is increased gradually and uniformly with time, which makes it easier to modulate the temperature over time. This provides an increased sensitivity as the separation range expands [37].

In DGGE/TGGE/TTGE, DNA is PCR-amplified using primers targeting part of the studied gene, mostly 16S rRNA gene. The 5'-end of the forward primer contains 35-40 base pair GC clamp which allows a certain portion of the DNA to remain double-stranded during electrophoresis. The denaturation step disrupts the hydrogen bonds holding together double strands of DNA [28], hence the separation in the gel occurs according to melting behavior of the double-stranded DNA which is proportional to G+C content and distribution [38]. Theoretically, DGGE/TGGE/TTGE can separate DNA with as little as one base pair difference. Again, each of these techniques have the limitation of PCR biases and a rather difficult sample handling process. In addition, complex communities such as soil are usually comprised of numerous populations in relatively equivalent proportions which can make it difficult to identify individual populations [28]. The advantages of these methods include reproducibility, rapidity, and ability to simultaneously analyze a large number of samples. A major advantage of these methods is that the bands can be sampled directly from the gel and

their nucleotide sequence determined, thus allowing the identification of putative numerically dominant bacteria [3,28].

(2.1.5) Single-strand conformation polymorphism (SSCP)

Like DGGE/TGGE, this method was originally developed to detect point mutations in DNA [39]. In SSCP, the separation of single-stranded DNA is performed on the polyacrylamide gel, and differences in mobility are caused by differences in the folded secondary structure. Because no denaturants are used, DNA molecules of equal size fold and hence move differently according to the DNA sequences. SSCP has similar advantages as DGGE/TGGE/TTGE as well as the same limitations. In addition, some single-stranded DNA can form more stable conformations and one sequence may be thus represented by more bands in the gel (for review see [1,12]).

(2.1.6) Ribosomal intergenic spacer analysis (RISA)

RISA is similar in principal to other fingerprinting techniques. Intergenic spacer region between 16S and 23S rRNA genes may encode tRNAs and, because of heterogeneity in its length and sequence, is useful for differentiating between bacterial strains and closely related species. In RISA, the intergenic spacer region is PCR-amplified, denatured, and separated on polyacrylamide gel. The sequence polymorphisms are detected by silver stain, or in the case of automated ribosomal intergenic spacer analysis (ARISA) the forward primer is fluorescently labeled and is automatically detected, which increases the sensitivity of the method. Although both methods are subject to the traditional limitations of PCR, they provide highly reproducible bacterial community profiles [40].

(2.2) GUANINE-CYTOSINE (G+C) CONTENT ANALYSIS

G+C content analysis is based on the concept that bacterial DNA varies in G+C content from 24% to 76% versus adenine-thymine (A+T). Particular taxonomic groups only include organisms that vary in G+C content by no more than 3-5%. Different taxonomic groups may share the same G+C range, and therefore, this method provides a coarse level of resolution, i.e. above genus level [41]. G+C content is usually determined by analyzing the melting curves of DNA since melting temperature of DNA is dependent on its G+C content. The melting profiles are converted to mol percent G+C profiles by calculating the first derivative of the melting curve [3].

G+C content analysis is not influenced by PCR biases, is quantitative, and is capable of revealing rare members of the community. However, large quantities of DNA are required which may limit the use of the method [41].

(2.3) REASSOCIATION KINETICS AND HYBRIDIZATION

The genetic complexity of the bacterial community can be determined by reassociation (reannealing) kinetics of sheared and thermally denatured (i.e. single-stranded) DNA. Under defined conditions the reassociation rate is proportional to the DNA complexity. The reassociation rate is determined relative to DNA with known complexity such as *Escherichia coli* DNA [3]. Alternatively, the similarities between communities of two different samples can be studied by measuring the degree of DNA similarity through hybridization kinetics [42].

Insight into microbial community structure is also provided by fluorescent *in situ* hybridization (FISH) as summarized by Wagner et al. (2003). FISH is based on rRNA targeted oligonucleotide probes, usually 15-25 nucleotides in length, labeled covalently at the 5'-end with a fluorescent dye. Bacterial cells hybridize under stringent conditions either in

solution or on a glass slide with these probes. Specifically stained cells are detected by flow cytometry or epifluorescence analysis [43].

(2.4) METAGENOMICS

Metagenomics (also ecological genomics, community genomics, or environmental genomics) is the discipline that uses genomics (functional analysis of entire genomes as well as mapping, sequencing, and analyzing genomes) methods to analyze natural ecological communities, namely the collective genomes in an environmental community [14]. Its power lies in sequencing as fully as possible the genomes of uncultured microbes, thus allowing the investigation of the broad diversity of either individual genes or whole operons [44]. Metagenomics comprises the extraction of the genomic DNA of a whole habitat, i.e. metagenomic DNA, and its further fragmentation and cloning, and subsequent analysis of constructed metagenomic library, or alternatively direct sequencing of metagenomic DNA. Although still in its infancy, the field of metagenomics has already discovered a richness of genomic and metabolic diversity that had not been previously imagined. Therefore, metagenomics is considered very promising in helping to answer some key questions in microbial ecology, such as how diverse metabolic pathways are, how microorganisms interact with each other, how many nutrient and energy sources are metabolized by different microorganisms and how many microorganisms and their enzymes metabolize these substrates, or how environmental stimuli impact ecosystem functioning [45].

Most metagenomic techniques had existed for some time before the first so called metagenomic studies, e.g. the phylogenetic analysis of environmental microbial diversity is an early form of metagenomics. However, several trends for metagenomic studies have emerged lately, including larger DNA fragments being cloned, increased number of sequences generated in individual studies, or increased number of niches being studied [14].

The large volume of sequencing was enabled by the next generation technology of 454 sequencing developed by 454 Life Sciences (now Roche). This technology has revolutionized pyrosequencing by higher throughput, simpler sample preparation, and miniaturization the sequencing reactions. Combination of these factors allows for parallel performing of many sequencing reactions at a time [46]. The 454 sequencing can also be used for amplicon sequencing. For bacterial diversity assessments, pyrosequencing of 16S rRNA gene amplicons is of special importance as it provides many times more sequence information compared to Sanger sequencing [47].

(2.3) MICROARRAYS

As previously discussed, microbial communities in contaminated soil systems can be incredibly heterogeneous and diverse. Accurately characterizing these environments is necessary for a full understanding of microbial community metabolic processes and functions. Genetic microarrays are a high-throughput method for detecting the presence and concentration of sequences within an environmental sample [48]. The technology works by spotting DNA oligonucleotide probes onto a matrix and each microarray can contain thousands of different probes [12]. These probes are designed to bind only complementary DNA, however, the stringency of nucleotide matching can be regulated by manipulating binding conditions, which can capture more diversity, but create the need for further downstream characterizations. Presence, and in some cases abundance, of target sequences on the array are indicated by fluorescence [49].

Microarrays were originally developed for monitoring whole-genome gene expressions. Recently, several types of microarrays have been introduced and evaluated for bacterial detection and community analyses [50]. These include phylogenetic oligonucleotide arrays, community genome arrays, and functional gene arrays. Phylogenetic oligonucleotide

arrays contain 18-25 bp of ribosomal rRNA genes. These probes can be designed to survey different levels of sequence conservation, from highly conserved sequences to variable sequences, thus enabling groupings at the level of species. Microarrays containing entire genomic DNAs are referred to as community genomic arrays. These arrays are a quantitative tool for the detection and identification of microorganisms in environmental samples. The third type of microarrays contain functional gene sequence information with 50 mer probes. These arrays are primarily used for functional analysis of microbial community activity in the environment [48].

Although traditional techniques such as PCR and cloning are necessary for deeper characterizations of microbial communities, they can be expensive, time consuming, introduce bias (due to primer design, etc.), and possibly not capture important members of the soil community due to the sheer complexity and heterogeneity of the system. Microarrays can avoid the culture bias associated with traditional cloning and sequencing approaches, and allow for the detection of even rare genes and taxa in the environment. By stringent probes design and due to the huge number of probes available on a microarray, they are able to create an accurate profile of communities. Microarrays can also help to link a large number of microbial identities and functions to geochemical parameters [49]. Microarrays have been widely used to study contaminated environments, both in combination with other molecular techniques and as a standalone method of community analysis [51,52,53,54].

Phylogenetic microarrays allow for the culture-independent identification of microorganisms in any type of sample [51]. These arrays can have probes for group and species level identification by either targeting the variable or conserved regions of the 16S rRNA gene [50]. Phylogenetic microarrays provide a comprehensive profile of the total microbial community and can be used as a filter to help direct sampling and experimental protocols.

Functional genetic microarrays (FGA) contain probes which target genes involved in microbially mediated biogeochemical and ecological processes in natural settings. Such processes can include organic contamination degradation, metal reduction and resistance, and C, N, P, and S utilization and cycling. One example of a widely used FGA is called the GeoChip which contains over 24,000 different oligonucleotide probes and covers more than 10,000 gene sequences [49]. The GeoChip has been used in a variety of ways to look at, for example, functional genes involved in PCB degradation [52], to monitor changes in microbial functional community throughout remediation of crude oil contaminated soils [53], and to monitor microbial communities in uranium contaminated groundwater [54].

(2.4) STABLE ISOTOPE PROBING

In contaminated environments, techniques for microbial diversity analysis which enable the linkage of metabolic process, in this instance pollutant metabolism, directly with specific members of the community are extremely important for understanding community dynamics. Stable isotope probing (SIP) is a technique that has been developed in recent years (for review see [55,56]) which allows one to probe an environmental sample with a stable isotope in order to discover which microbes are performing a specific role in the environment. This is done usually through incubating a sample with a compound of interest which has been labeled with a heavier isotope of carbon, nitrogen or oxygen. This insertion of the labeled compound allows the microbes to utilize it as if in nature, after which DNA [57], RNA [58], phospholipid fatty acids [59], or proteins [60] can be extracted and analyzed. Specifically in the case of DNA-based stable isotope probing (DNA-SIP) a ^{13}C-labeled substrate is provided to a microbial community and subsequent analysis of the DNA isolated from the community is performed. Isopycnic centrifugation permits separating ^{13}C-labeled DNA of organisms that utilized the substrate from ^{12}C-DNA of the inactive majority. As the whole metagenome of

active populations is isolated, its follow-up analysis provides successful taxonomic identification as well as the potential for functional gene analyses. Because of its power, SIP has become one of the leading techniques of microbial ecology research. Examples of SIP are wide spread and this highly useful technique has been used to investigate methanotrophs and methylotrophs in different habitats [57,61,62,63,64,65,66], to examine ammonia-oxidizing sediment bacteria [67], to identify bacteria involved in degradation of glucose by anoxic slurries of a forest soil [68], to analyze the microbial community of a Douglas-fir forest [69], to identify cellulolytic bacteria [70] as well as to determine bacteria associated with the metabolism of (polychlorinated) biphenyls [52,71,72], pentachlorophenol [73], or methyl chloride [74].

(3) Fatty acid methyl ester (FAME) analysis

Fatty acid methyl ester analysis is a biochemical method that does not rely on culturing microorganisms. Cell biomass is made up of a relatively constant proportion of fatty acids, and major taxonomic groups within a community can be differentiated through analysis of signature fatty acids. Therefore, their analysis can provide insight to the microbial community and composition of viable microbial taxa in the experimental system [75].

Upon extracting fatty acids directly from the soil, methylation is performed, and FAMEs are analyzed by gas chromatography . Since FAME analysis relies on extracting fatty acids directly from soil, it overcomes the limitations of culture-dependent methods. However, the fatty acid composition can be influenced by external factors, and the profiles can be confounded by other organisms [1].

Estimations and measures of microbial diversity

Although all broad methodologies described above provide information about microbial diversity, it is especially 16S rRNA genes that afford both recognition and classification of the prokaryotic hosts of the genes [76]. Therefore, estimations and measures of microbial diversity are commonly based on these genes. Microbial diversity within a habitat, community, or ecosystem is usually expressed by rarefaction curve approach and/or by diversity indices.

A rarefaction curve approach measures the accumulation of distinctive taxa as a function of sampling effort. This approach presumes that, as sampling proceeds, members of the rarest species (or OTUs) will gradually be added to the total. Therefore, the asymptote in the rarefaction accumulation curve estimates total diversity in the habitat sampled [76].

A nonparametric estimation of the microbial diversity, the Chao1 estimator, documents the absolute number of species or OTUs in an assemblage [76,77]. It is calculated as a proportion of species (OTUs) that have been recaptured relative to those that have been captured only once [78]:

$$S_{\text{Chao1}} = \text{S}_{\text{obs}} + \frac{n_1\left(n_1 - 1\right)}{2\left(n_2 + 1\right)} \qquad \text{(Formula 1)}$$

where S_{obs} is the number of observed species (OTUs), n_1 is the number of species (OTUs) captured once (singletons), and n_2 is the number of species (OTUs) captured twice (doubletons).

The Shannon index takes into account the species (or OTU) richness and evenness and is increased when either the richness or evenness is increased [77,78]:

$$H_{\text{Shannon}} = -\sum_{i=1}^{\text{S}_{\text{obs}}} \frac{n_i}{N} \ln \frac{n_i}{N} \qquad \text{(Formula 2)}$$

where S_{obs} is the number of observed species (OTUs), n_i is the number of species (OTUs) with *i* individuals, and N is the total number of individuals.

The Simpson index is insensitive to the taxa richness and represents the probability that any two randomly selected individuals belong to the same species or OTU [77]:

$$D_{\text{Simpson}} = \frac{\sum_{i=1}^{S_{obs}} n_i(n_i - 1)}{N(N-1)} \qquad \text{(Formula 3)}$$

where S_{obs} is the number of observed species (OTUs), n_i is the number of species (OTUs) with *i* individuals, and N is the total number of individuals.

LIST OF SELECTED WEB PAGES IMPORTANT FOR MICROBIAL ECOLOGY AND BIOREMEDIATION RESEARCH

Ribosomal Database Project (https://rdp.cme.msu.edu/index.jsp)

Ribosomal Database Project (RDP) provides the research community with aligned and annotated rRNA gene sequences, along with analysis services and a phylogenetically consistent taxonomic framework for these data. RDP release 10, update 18 (January 2010) contains 1,358,426 16S rRNA gene sequences in aligned and annotated format [79,80,81].

Functional Gene Pipeline/Repository (http://fungene.cme.msu.edu)

Functional Gene Pipeline/Repository (FGPR) is an interactive display of sequence search results for those interested in a particular gene family as well as a tool to aid functional genomics studies, especially of the environment.

FastGroupII (http://biome.sdsu.edu/fastgroup/fg_tools.htm)

FastGroupII is a tool to investigate the biodiversity of microbial communities. It trims the DNA/RNA sequences with the user specified parameters, dereplicates them, and groups the similar sequences together [82].

MEGA 4: Molecular Evolutionary Genetics Analysis (http://www.megasoftware.net)

MEGA is an integrated tool for conducting automatic and manual sequence alignment, inferring phylogenetic trees, mining web-based databases, estimating rates of molecular evolution, and testing evolutionary hypotheses [83].

MOTHUR (http://schloss.micro.umass.edu)

MOTHUR enables investigators to align their 16S rRNA gene sequences, calculate pairwise distances, and analyze the resulting distance matrices using the algorithms employed in LIBSHUFF / ∫−LIBSHUFF, DOTUR, SONS, TreeClimber, and UniFrac using a single program.

MiCA: Microbial Community Analysis (http://mica.ibest.uidaho.edu)

MiCA facilitates analyses of terminal-restriction fragment length polymorphisms (T-RFLP) of 16S and 18S rRNA genes by performing virtual polymerase chain reaction (PCR) amplification of rRNA genes and restriction of the amplicons using primer sequences and restriction enzymes chosen by the user and by inferring a plausible community structure(s) based on T-RFLP data supplied by a user [84].

CERCLA Priority List of Hazardous Substances (http://www.atsdr.cdc.gov/cercla)

CERCLA (Comprehensive Environmental Response, Compensation, and Liability Act) Priority List of Hazardous Substances includes the substances that are determined to pose the

most significant potential threat to human health due to their known or suspected toxicity. This CERCLA priority list is revised and published on a 2-year basis, with a yearly informal review and revision.

CONCLUDING REMARKS

Over the last two decades, the knowledge of bacterial biodiversity has expanded due to sequence analyses of 16S rRNA genes amplified and cloned from DNA isolated directly from the environment. These approaches have demonstrated that bacterial diversity is far more extensive than ever imagined from culture-based studies. More recently, surveys using 16S rRNA gene have been eclipsed by metagenomics, the study of collective genomes in the community, as it can provide a global overview of a community's functional potential rather than solely its phylogenetic composition [13,14]. However, neither 16S rRNA gene analyses nor metagenomics are exempt from limitations; PCR biases cannot be ruled out when 16S rRNA gene analyses are performed, whereas metagenomic analyses are likely to be misinterpreted by inappropriate predictions caused by annotation errors in databases. To overcome this, genomes of many more cultivated bacteria should be investigated to provide valuable references for metagenomic annotation and prediction, which highlights the importance of culture-based studies. These still have not been overcome as a means to associate functions to newly discovered microbes and to test physiologically potential functions eventually deduced from gene amplification or metagenomic data [5]. Due to advantages and limitations and unique insights provided by culture-based and molecular techniques, only the combination of these approaches can currently provide an accurate view of the community.

REFERENCES

1. Zhang, L., and Xu, Z. 2008, J. Soils Sediments, 8, 379.
2. Torsvik, V., Ovreas, L., and Thingstad, T.F. 2002, Science, 296, 1064.
3. Torsvik, V., Daae, F.L., Sandaa, R.A., and Ovreas, L. 1998, J. Biotechnol., 64, 53.
4. Lozupone, C.A., and Knight, R. 2007, Proc. Natl. Acad. Sci. U.S.A., 104, 11436.
5. Lopez-Garcia, P., and Moreira, D. 2008, Res. Microbiol., 159, 67.
6. Stursa, P., Uhlik, O., Kurzawova, V., Koubek, J., Ionescu, M., Strohalm, M., Lovecka, P., Macek, T., and Mackova, M. 2009, Plant Soil Environ., 55, 389.
7. Torsvik, V., and Ovreas, L. 2002, Curr. Opin. Microbiol., 5, 240.
8. Lozupone, C.A., and Knight, R. 2008, FEMS Microbiol. Rev., 32, 557.
9. Macek, T., Mackova, M., and Kas, J. 2000, Biotechnol. Adv., 18, 23.
10. Macek, T., Uhlik, O., Jecna, K., Novakova, M., Lovecka, P., Rezek, J., Dudkova, V., Stursa, P., Vrchotova, B., Pavlikova, D., Demnerova, K., and Mackova, M. 2009, Advances in phytoremediation and rhizoremediation, A. Singh, R.C. Kuhad, and O.P. Ward (Ed.), Springer, Berlin, Germany, 257.
11. Baker, G.C., Smith, J.J., and Cowan, D.A. 2003, J. Microbiol. Methods, 55, 541.
12. Kirk, J.L., Beaudette, L.A., Hart, M., Moutoglis, P., Klironomos, J.N., Lee, H., and Trevors, J.T. 2004, J. Microbiol. Methods, 58, 169.
13. Tringe, S.G., and Hugenholtz, P. 2008, Curr. Opin. Microbiol., 11, 442.
14. Xu, J. 2006, Mol. Ecol., 15, 1713.
15. Achtman, M., and Wagner, M. 2008, Nat. Rev. Microbiol., 6, 431.
16. Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T., Feil, E.J., Stackebrandt, E., Van de Peer Y., Vandamme, P., Thompson, F.L., and Swings, J. 2005, Nat. Rev. Microbiol., 3, 733.
17. Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., and Tiedje, J.M. 2007, Int. J. Syst. Evol. Microbiol., 57, 81.
18. Cardenas, E., and Tiedje, J.M. 2008, Curr. Opin. Biotechnol., 19, 544.
19. Hanage, W.P., Fraser, C., and Spratt, B.G. 2006, Philos. Trans. R. Soc. Lond B Biol. Sci., 361, 1917.
20. Leigh, M.B. 2006, Methods for rhizoremediation research. Approaches to experimental design and microbial analysis, M. Mackova, D. Dowling, and T. Macek (Ed.), Springer, Dordrecht, Netherlands, 33.

21. Sutherland, J.B., and Rafii, F. 2006, Cultural, Serological, and Genetic Methods for Identification of Bacteria, C.L. Wilkins and J.O. Lay, Jr. (Ed.), John Wiley & Sons, Inc., Hoboken, NJ, USA, 1.
22. Liyanage, R., and Lay, J.O., Jr. 2006, An Introduction to MALDI-TOF MS, C.L. Wilkins and J.O. Lay, Jr. (Ed.), John Wiley & Sons, Inc., Hoboken, NJ, USA, 39.
23. Krieg, R., Maier, T., Pusch, W., Mix, G., Schwarz, G., Macht, M., and Kostrzewa, M. 2008, 12th International Meeting of the Association of Biomolecular Resource Facilities, Salt Lake City, Utah, 9. - 12. 2. 2008.
24. Maier, T., Schwarz, G., and Kostrzewa, M. 2005, http://www.bdal.de/uploads/media/MT-80-3-ident-eBook_01.pdf.
25. Sylvestre, M. 1980, Appl. Environ. Microbiol., 39, 1223.
26. Pace, N.R., Stahl, D.A., Lane, D.J., and Olsen, G.J. 1985, ASM News, 51, 4.
27. Mullis, K.B., and Faloona, F.A. 1987, Methods Enzymol., 155, 335.
28. Marsh, T.L., and Nakatsu, C.H. 2007, Analysis of Microbial Communities with Denaturing Gradient Gel Electrophoresis and Terminal Restriction Fragment Length Polymorphism, C.A. Reddy, T.J. Beveridge, J.A. Breznak, G. Marzluf, T.M. Schmidt, and L.R. Snyder (Ed.), ASM Press, Washington, DC, USA, 909.
29. Isenbarger, T.A., Finney, M., Rios-Velazquez, C., Handelsman, J., and Ruvkun, G. 2008, Appl. Environ. Microbiol., 74, 840.
30. Yokouchi, H., Fukuoka, Y., Mukoyama, D., Calugay, R., Takeyama, H., and Matsunaga, T. 2006, Environ. Microbiol., 8, 1155.
31. Arriola, E., Lambros, M.B., Jones, C., Dexter, T., Mackay, A., Tan, D.S., Tamber, N., Fenwick, K., Ashworth, A., Dowsett, M., and Reis-Filho, J.S. 2007, Lab Invest, 87, 75.
32. Rudi, K., Nogva, H.K., Moen, B., Nissen, H., Bredholt, S., Moretro, T., Naterstad, K., and Holck, A. 2002, Int. J. Food Microbiol., 78, 171.
33. Cupples, A.M. 2008, J. Microbiol. Methods, 72, 1.
34. Liu, W.T., Marsh, T.L., Cheng, H., and Forney, L.J. 1997, Appl. Environ. Microbiol., 63, 4516.
35. Marsh, T.L. 2005, Methods Enzymol., 397, 308.
36. Muyzer, G., de Waal, E.C., and Uitterlinden, A.G. 1993, Appl. Environ. Microbiol., 59, 695.
37. Vasquez, A., Ahrne, S., Pettersson, B., and Molin, G. 2001, Lett. Appl. Microbiol., 32, 215.

38. Fromin, N., Hamelin, J., Tarnawski, S., Roesti, D., Jourdain-Miserez, K., Forestier, N., Teyssier-Cuvelle, S., Gillet, F., Aragno, M., and Rossi, P. 2002, Environ. Microbiol., 4, 634.
39. Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. 1989, Genomics, 5, 874.
40. Fisher, M.M., and Triplett, E.W. 1999, Appl. Environ. Microbiol., 65, 4630.
41. Tiedje, J.M., Asuming-Brempong, S., Nusslein, K., Marsh, T.L., and Flynn, S.J. 1999, Appl. Soil Ecol., 13, 109.
42. Griffiths, B.S., Ritz, K., Ebblewhite, N., and Dobson, G. 1999, Soil Biol. Biochem., 31, 145.
43. Wagner, M., Horn, M., and Daims, H. 2003, Curr. Opin. Microbiol., 6, 302.
44. Schmeisser, C., Steele, H., and Streit, W.R. 2007, Appl. Microbiol. Biotechnol., 75, 955.
45. Vieites, J.M., Guazzaroni, M.E., Beloqui, A., Golyshin, P.N., and Ferrer, M. 2009, FEMS Microbiol. Rev., 33, 236.
46. Rothberg, J.M., and Leamon, J.H. 2008, Nat. Biotechnol., 26, 1117.
47. Engelbrektson, A., Kunin, V., Wrighton, K.C., Zvenigorodsky, N., Chen, F., Ochman, H., and Hugenholtz, P. 2010, ISME J., *in press*, doi:10.1038/ismej.2009.153.
48. Zhou, J. 2003, Curr. Opin. Microbiol., 6, 288.
49. He, Z., Gentry, T.J., Schadt, C.W., Wu, L., Liebich, J., Chong, S.C., Huang, Z., Wu, W., Gu, B., Jardine, P., Criddle, C., and Zhou, J. 2007, ISME J., 1, 67.
50. Andersen, G.L., He, Z., DeSantis, T.Z., Brodie, E.L., and Zhou, J. 2010, The Use of Microarrays in Microbial Ecology, W.T. Liu and J.K.Jansson (Ed.), Caister Academic Press, Norfolk, UK, 87.
51. Brodie, E.L., DeSantis, T.Z., Parker, J.P., Zubietta, I.X., Piceno, Y.M., and Andersen, G.L. 2007, Proc. Natl. Acad. Sci. U.S.A., 104, 299.
52. Leigh, M.B., Pellizari, V.H., Uhlik, O., Sutka, R., Rodrigues, J., Ostrom, N.E., Zhou, J., and Tiedje, J.M. 2007, ISME J., 1, 134.
53. Liang, Y., Nostrand, J.D., Wang, J., Zhang, X., Zhou, J., and Li, G. 2009, Chemosphere, 75, 193.
54. Van Nostrand, J.D., Wu, W.M., Wu, L., Deng, Y., Carley, J., Carroll, S., He, Z., Gu, B., Luo, J., Criddle, C.S., Watson, D.B., Jardine, P.M., Marsh, T.L., Tiedje, J.M., Hazen, T.C., and Zhou, J. 2009, Environ. Microbiol., 11, 2611.
55. Uhlik, O., Jecna, K., Leigh, M.B., Mackova, M., and Macek, T. 2009, Sci. Total Environ., 407, 3611.

56. Uhlik, O., Jecna, K., Mackova, M., Leigh, M.B., Demnerova, K., and Macek, T. 2008, Chem. Listy, 102, 474.
57. Radajewski, S., Ineson, P., Parekh, N.R., and Murrell, J.C. 2000, Nature, 403, 646.
58. Manefield, M., Whiteley, A.S., Griffiths, R.I., and Bailey, M.J. 2002, Appl. Environ. Microbiol., 68, 5367.
59. Boschker, H.T.S., Nold, S.C., Wellsbury, P., Bos, D., de Graaf, W., Pel, R., Parkes, R.J., and Cappenberg, T.E. 1998, Nature, 392, 801.
60. Jehmlich, N., Schmidt, F., Hartwich, M., von, B.M., Richnow, H.H., and Vogt, C. 2008, Rapid Commun. Mass Spectrom., 22, 2889.
61. Bengtson, P., Basiliko, N., Dumont, M.G., Hills, M., Murrell, J.C., Roy, R., and Grayston, S.J. 2009, FEMS Microbiol. Ecol.,
62. Han, B., Chen, Y., Abell, G., Jiang, H., Bodrossy, L., Zhao, J., Murrell, J.C., and Xing, X.H. 2009, FEMS Microbiol. Ecol., 70, 40.
63. Jensen, S., Neufeld, J.D., Birkeland, N.K., Hovland, M., and Murrell, J.C. 2008, FEMS Microbiol. Ecol., 66, 320.
64. Chen, Y., Dumont, M.G., Neufeld, J.D., Bodrossy, L., Stralis-Pavese, N., McNamara, N.P., Ostle, N., Briones, M.J., and Murrell, J.C. 2008, Environ. Microbiol., 10, 2609.
65. Chen, Y., Dumont, M.G., McNamara, N.P., Chamberlain, P.M., Bodrossy, L., Stralis-Pavese, N., and Murrell, J.C. 2008, Environ. Microbiol., 10, 446.
66. Neufeld, J.D., Chen, Y., Dumont, M.G., and Murrell, J.C. 2008, Environ. Microbiol., 10, 1526.
67. Freitag, T.E., Chang, L., and Prosser, J.I. 2006, Environ. Microbiol., 8, 684.
68. Degelmann, D.M., Kolb, S., Dumont, M., Murrell, J.C., and Drake, H.L. 2009, FEMS Microbiol. Ecol., 68, 312.
69. Moore-Kucera, J., and Dick, R.P. 2008, Microb. Ecol., 55, 500.
70. Haichar, F.Z., Achouak, W., Christen, R., Heulin, T., Marol, C., Marais, M.F., Mougel, C., Ranjard, L., Balesdent, J., and Berge, O. 2007, Environ. Microbiol., 9, 625.
71. Uhlik, O., Jecna, K., Mackova, M., Vlcek, C., Hroudova, M., Demnerova, K., Paces, V., and Macek, T. 2009, Appl. Environ. Microbiol., 75, 6471.
72. Sul, W.J., Park, J., Quensen, J.F., III, Rodrigues, J.L., Seliger, L., Tsoi, T.V., Zylstra, G.J., and Tiedje, J.M. 2009, Appl. Environ. Microbiol., 75, 5501.
73. Mahmood, S., Paton, G.I., and Prosser, J.I. 2005, Environ. Microbiol., 7, 1349.

74. Borodina, E., Cox, M.J., McDonald, I.R., and Murrell, J.C. 2005, Environ. Microbiol., 7, 1318.
75. Ibekwe, A.M., and Kennedy, A.C. 1998, FEMS Microbiol. Ecol., 26, 151.
76. Madsen, E.L. 2008, Environmental microbiology. From genomes to biogeochemistry, Blackwell Publishing, Malden, MA, USA.
77. Magurran, A.E. 2004, Measuring biological diversity, Blackwell Publishing, Malden, MA, USA.
78. Schloss, P.D., Westcott, S., and Ryabin, T. 2009, http://schloss.micro.umass.edu/mothur.
79. Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R.J., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., Marsh, T., Garrity, G.M., and Tiedje, J.M. 2009, Nucleic Acids Res., 37, D141.
80. Cole, J.R., Chai, B., Farris, R.J., Wang, Q., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., Bandela, A.M., Cardenas, E., Garrity, G.M., and Tiedje, J.M. 2007, Nucleic Acids Res., 35, D169.
81. Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. 2007, Appl. Environ. Microbiol., 73, 5261.
82. Yu, Y., Breitbart, M., McNairnie, P., and Rohwer, F. 2006, BMC Bioinformatics, 7, 57.
83. Tamura, K., Dudley, J., Nei, M., and Kumar, S. 2007, Mol. Biol. Evol., 24, 1596.
84. Shyu, C., Soule, T., Bent, S.J., Foster, J.A., and Forney, L.J. 2007, Microb. Ecol., 53, 562.

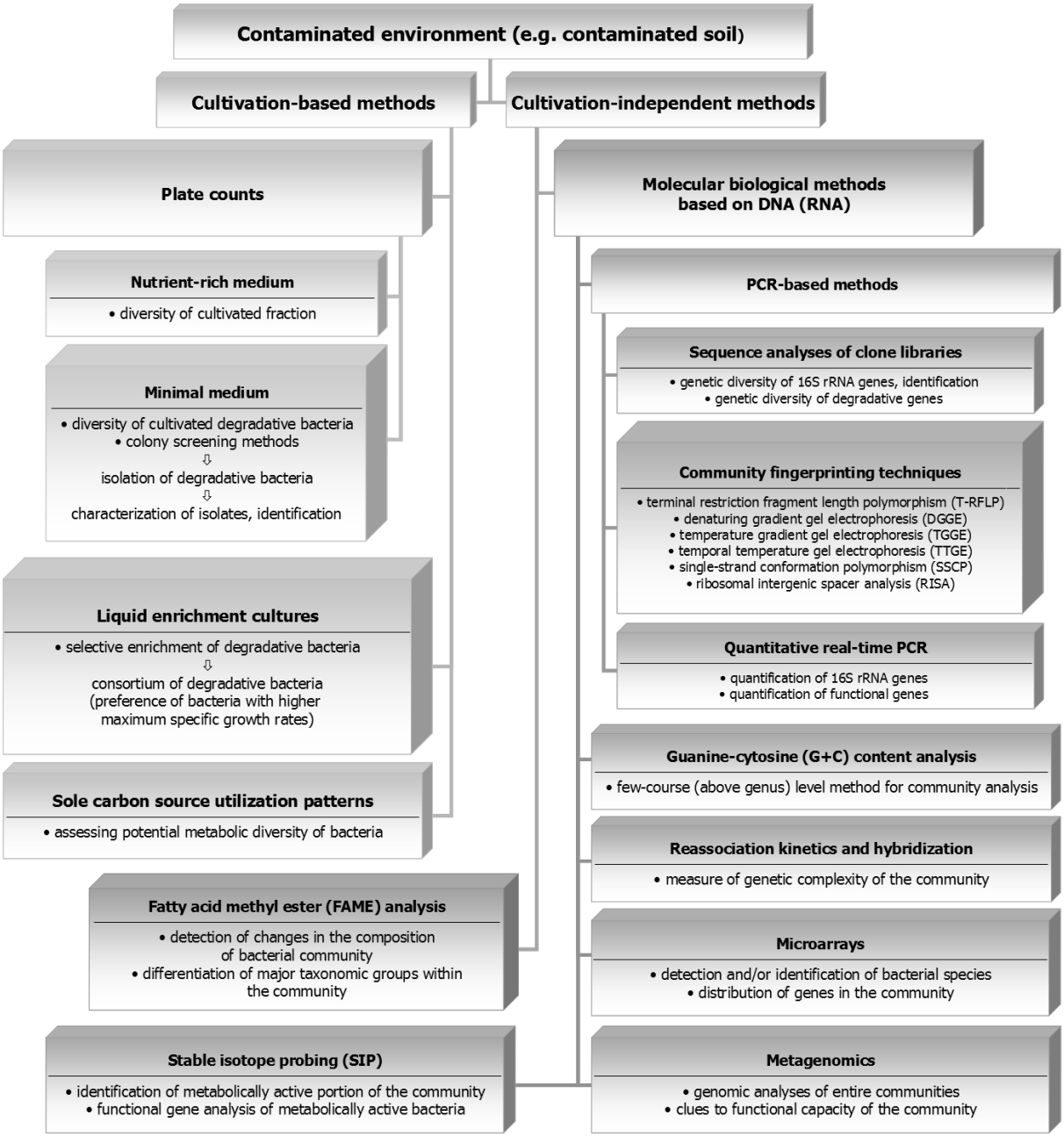

Fig. A1-1: Overview of methods commonly used for bacterial population analyses in contaminated soil.

Appendix 2

Stable Isotope Probing in the Metagenomics Era: A Bridge Towards Improved Bioremediation [1]

ABSTRACT

Microbial biodegradation and biotransformation reactions are essential to most bioremediation processes, yet the specific organisms, genes, and mechanisms involved are often not well understood. Stable isotope probing (SIP) enables researchers to directly link microbial metabolic capability to phylogenetic and metagenomic information within a community context by tracking isotopically labeled substances into phylogenetically and functionally informative biomarkers. SIP is thus applicable as a tool for the identification of active members of the microbial community and associated genes integral to the community functional potential, such as biodegradative processes. The rapid evolution of SIP over the last decade and integration with metagenomics provides researchers with a much deeper insight into potential biodegradative genes, processes, and applications, thereby enabling an improved mechanistic understanding that can facilitate advances in the field of bioremediation.

[1] Uhlik, O., Leewis, M.C., Strejcek, M., Musilova, L., Leigh, M.B., Macek, T. 2013. Stable isotope probing in the metagenomics era: a bridge towards improved bioremediation. Biotechnology Advances. 31(2): 154-165. http://dx.doi.org/10.1016/j.biotechadv.2012.09.003.
Note: Leewis and Uhlik contributed equally to this work.

INTRODUCTION

In addition to their critical role in global biogeochemical cycling, microbes play an essential role in the degradation, mineralization, or transformation of environmental pollutants and thus are potentially capable of restoring contaminated sites (Liu and Suflita, 1993, Diaz, 2004, Geoffrey, 2010). The natural biodegradative processes occurring in contaminated sites are known as intrinsic bioremediation or natural attenuation (Mulligan and Yong, 2004, Jørgensen, 2007). These microbial biodegradative processes can often be accelerated using various strategies, known as bioremediation. Together with phytoremediation, microbially mediated bioremediation is generally considered an environmentally friendly, inexpensive, publicly accepted and promising means to remove contaminants from the environment (Schnoor et al., 1995, Macek et al., 2000).

When designing a strategy for the bioremediation of any contaminated site, an understanding of the indigenous microbial community can be highly valuable. Information integral to the success of bioremediation strategies can include: (i) identification of the microorganisms present in the contaminated site, (ii) investigation of their metabolic capabilities, and (iii) understanding of potential microbial community shifts in response to changing environmental factors (Lovley, 2003). Much of this information, however, has been difficult to elucidate and is rarely achieved with traditional microbiological techniques. Until recently, knowledge of microbes involved in bioremediation processes had been based mostly on culture-dependent studies, which do not account for the fact that laboratory conditions differ substantially from the environment. Cultivation-based techniques have been shown to target only about 1% of microbes occurring in the environment (Torsvik and Øvreås, 2002, Lozupone and Knight, 2008, Zhang and Xu, 2008). More importantly, cultivation often fails to predict which microbes and which specific metabolic pathways will be active under realistic environmental conditions (Morales and Holben, 2011). The advent of molecular microbial ecology enabled

culture-independent phylogenetic analyses of communities and functional genes, however, linking contaminant transformation to phylogenetic identity and specific genes/enzymes of metabolically active microbes without cultivation remained a major challenge and still required cultivation. The development of stable isotope probing (SIP) was instrumental in circumventing the limitations of culture based investigations of biodegradation (Wellington et al., 2003, Friedrich, 2006). The coupling of SIP with the rapidly advancing field of metagenomics is further increasing its potential benefits to the field of bioremediation. This review aims to discuss the emerging trends in stable isotope probing techniques integrated with metagenomics, which can provide researchers with an unparalleled understanding of contaminant biodegradation and biotransformation processes in the environment.

PRINCIPLES OF STABLE ISOTOPE PROBING (SIP)

SIP tracks the incorporation of heavy stable isotopes, mainly ^{13}C (Radajewski et al., 2003, Dumont and Murrell, 2005, Neufeld et al., 2007b, Uhlík et al., 2009a, Chen and Murrell, 2010, Madsen, 2010), ^{15}N (Buckley et al., 2007b, Buckley et al., 2007a, Roh et al., 2009, Bell et al., 2011), or rarely ^{18}O and ^{2}H (Aanderud and Lennon, 2011, Woods et al., 2011), from specific substrates into phylogenetically informative biomarkers associated with microbes that assimilate the substrate. After stable isotopes have been pulsed into the environment and metabolically active cells have incorporated the label into their biomass, biomarkers are recovered and analyzed (Neufeld et al., 2007b, Dunford and Neufeld, 2010) (Figure A2-1). Therefore, SIP is an approach that can identify microbial populations with a defined function. The first biomarkers introduced in SIP studies were phospholipid-derived fatty acids (PLFA) (Boschker et al., 1998), followed by DNA (Radajewski et al., 2000), and rRNA (Manefield et al., 2002). Although it is the most sensitive of the three, PLFA-SIP is restricted to the classification of broad taxonomic groups (Table A2-1). Analyzing nucleic acids (RNA or

DNA) can be far more informative taxonomically, with rRNA being a more responsive biomarker than DNA. This is due to the fact that the rates of rRNA synthesis are much higher than those of DNA, and rRNA labeling is independent of cellular replication (Manefield et al., 2002, Manefield et al., 2004, Whiteley et al., 2006). In addition, reverse transcription and sequence analysis of 16S rRNA provides equal resolution for phylogenetic identification as sequence analysis of 16S rRNA genes (DNA, Table A2-1). Ribosomal RNA-based analyses, however, cannot provide access to the functional genes responsible for the metabolic capabilities of the community, and generally ^{13}C-labeled mRNA is challenging to isolate in sufficient quantities for SIP (Neufeld et al., 2007a, Simon and Daniel, 2009). However, success has been achieved with mRNA-SIP for detection of naphthalene dioxygenase transcripts following SIP incubations with naphthalene (Huang et al., 2009). Further research employing DNA-, mRNA-, and rRNA-SIP (Dumont et al., 2011) demonstrated higher rates of labeling of functional mRNAs than their genes, and hence higher sensitivity of mRNA-SIP compared to DNA-SIP (Table A2-1). In addition, protein-SIP has recently been developed, taking advantage of proteins as a combined indicator for a specific metabolic activity as well as for obtaining phylogenetic information (Table A2-1, Jehmlich et al., 2008a, Jehmlich et al., 2008b, Jehmlich et al., 2009).

To date, when functional gene analyses coupled to specific assimilatory processes are desired, DNA-SIP has been most often employed, due to the relative ease of detecting functional genes in ^{13}C-DNA compared to mRNA. The use of either DNA-SIP or RNA-SIP can enable both the phylogenetic identification of those microbes performing the process as well as key metabolic genes possessed by the active populations likely to be involved in the process. When coupled with the rapidly expanding field of metagenomics, DNA-SIP has the potential for a focused, in-depth analysis of the collective genomes of the community active in a particular biodegradative process. This provides definitive information about the

populations active in specific assimilatory processes and also requires considerably less sequencing effort than a full metagenomic analysis of the total community.

THE METAGENOMICS ERA

Metagenomics (also known as ecological genomics, community genomics, or environmental genomics) is a discipline that uses genomic methods to analyze natural ecological communities, namely the collective genomes in an environmental community (Handelsman et al., 1998, Riesenfeld et al., 2004). The major goal of metagenomics is to explicate the genomes of uncultured microbes, thereby permitting investigation of the broad diversity of taxonomically and phylogenetically relevant genes, individual catabolic genes, and whole operons (Schmeisser et al., 2007). Metagenomics itself was initially recognized for its potential to aid in discovery of novel biomolecules for biotechnological applications (Riesenfeld et al., 2004). Although the basic concept of metagenomics was first introduced at the end of last century (Handelsman et al., 1998), early forms of metagenomics had begun to emerge previously, with one example being the phylogenetic analyses of environmental microbial communities (Pace et al., 1985).

The approach introduced by Handelsman et al. (1998) involves extraction of the metagenome (genomic DNA from all organisms inhabiting the environment), its fragmentation, cloning, transformation, and subsequent screening of the constructed metagenomic library. The primary aim is to screen environmental communities for a specific biological activity and identify genes or gene clusters associated with it; also referred to as function-based screening (Yun and Ryu, 2005). The advent of high-throughput next generation DNA sequencing (e.g. 454 pyrosequencing, Illumina (Solexa) sequencing, SOLiD sequencing) gave rise to another approach to metagenomics, sequence-based screening. This method was first demonstrated by environmental genome shotgun sequencing of the Sargasso Sea (Venter et al., 2004),

showing the potential of revealing vast phylogenetic and metabolic diversity of microbial communities (Yun and Ryu, 2005). In addition to sequence-based screening of environmental metagenomic libraries, direct pyrosequencing of environmental communities is possible, bypassing cloning completely (Edwards et al., 2006). Because communities from habitats such as soils and sediments, are often too diverse to permit screening in sufficient depth, even with the use of high-throughput sequencing, gene-targeted metagenomics techniques have emerged (Iwai et al., 2011), which are based on sequencing of PCR-generated amplicons (Figure A2-2).

In addition to high-throughput sequencing, metagenomic analyses have recently been performed with the use of high-throughput microarrays (Figure A2-2). These have been used to analyze microbial communities and monitor environmental biogeochemical processes. GeoChip microarrays (He et al., 2007, He et al., 2010) currently contain 83,992 50-meric sequences covering approximately 152,414 genes encoding for enzymes responsible for biogeochemical (C, N, P, S) cycling, metabolic processes, heavy metal resistance, antibiotic resistance, degradation of pollutants, and *gyrB* genes (Hazen et al., 2010, Lu et al., 2012). Marker *gyrB* encoding for the gyrase β-subunit is used instead of the more common 16S rRNA genes as probes for 16S rRNA usually do not provide resolution below genus level. *gyrB* can be used to differentiate even closely related species (He et al., 2010). GeoChip microarrays can be therefore used to study structure, dynamics, and potential metabolic activity of microbial communities and their variations depending on different stimuli. Another type of microarray valuable to microbial ecology and contaminant biodegradation is the PhyloChip, which is used for high throughput phylogenetic analyses of microbial communities (Brodie et al., 2006), and has been used for a variety of applications including assessing microbial community responses to petroleum contamination (Hazen et al., 2010, DeAngelis et al., 2011).

With the first applications of metagenomic techniques, it became apparent that they enable the discovery of genomic and metabolic diversity that had not been previously imagined (Schloss and Handelsman, 2005). As research progressed, however, the main drawbacks of metagenomics were realized: the inability to link specific functions to individual populations and to achieve full sequence coverage in more complex communities (Vieites et al., 2009). By combining metagenomic techniques with SIP, these limitations can be significantly reduced. SIP experiments are designed to provide a targeted analysis of the active populations, omitting the inactive majority that is not the focus of the study (Figure A2-2).

INTEGRATION OF SIP WITH METAGENOMICS

One of the main drawbacks of total community metagenomics is the unlikelihood of detecting particular genes of interest due to the extremely vast diversity and abundance of microbial genes occurring in most ecosystems, even when function-based metagenomic screening is used. Targeting metagenomics to specific subpopulations, as with SIP, which are likely to contain the genes of interest may overcome this obstacle (Schloss and Handelsman, 2003). This was demonstrated by Schwarz and colleagues (2006) who isolated genes encoding for coenzyme B_{12}-dependent glycerol dehydratases. The source for this key enzyme for the anaerobic dehydration of glycerol was the enrichment cultures of ^{13}C-glycerol-fermenting microorganisms from a sediment sample. When metagenomic library construction was preceded by SIP, the frequency of clones bearing target genes was increased by almost four fold. However, expressing target genes successfully *in vitro*, such as in metagenomic libraries, can be very challenging in some instances. Using high-throughput sequencing technologies for direct shotgun sequencing of SIP-derived metagenomes can aid in overcoming this obstacle. Selective enrichment of targeted populations, whose diversity is much less complex than that of total communities, with subsequent isolation of the particular

functional metagenome of interest increases the feasibility of achieving coverage and assembly of individual genomes with significantly reduced efforts and sequencing cost. Shotgun sequencing of SIP-derived metagenomes can also help ensure that portions of the microbial community that have low abundance but are integral to the metabolic processes of interest will not be overlooked (Wellington et al., 2003). The major drawback of this approach is usually the recovery of DNA in quantities too small to be sufficient for shotgun sequencing. Advances in multiple displacement amplification over the last years (Binga et al., 2008) have helped to overcome this limitation. Although no bioremediation studies have yet been published performing direct shotgun sequencing of SIP-derived metagenomes, they are very likely to arise in the near future.

SIP and metagenomics to study biodegradation of ecologically significant C_1 compounds

One of the pilot studies integrating DNA-SIP with metagenomics revealed a complete methane monooxygenase operon in forest soils (Dumont et al., 2006). Methane and other one-carbon (C_1) compounds are of global ecological significance because they can affect global warming, influence atmospheric and marine chemistry, and impact cloud formation. In the context of bioremediation, methanotrophs and/or methylotrophs have been implicated in the degradation of trichloroethylene and *cis*-1,2-dichloroethylene (Little et al., 1988, Arai et al., 1999, Shigematsu et al., 1999, Takeuchi et al., 2005), insecticides (Topp et al., 1993), nitro-substituted explosives (Van Aken et al., 2004), methyl halides (Warner et al., 2005), methyl *tert*-butyl ether (Nakatsu et al., 2006, Kane et al., 2007), and other xenobiotic compounds. Therefore, investigating the metabolism of C_1 compounds is also potentially valuable for bioremediation.

Dumont and colleagues (2006) were the first one to apply SIP in combination with function- and sequence-based metagenomic library screening in this field. After the incubation of a soil sample with $^{13}CH_4$, ^{13}C-DNA was used for the construction of a metagenomic library using a

bacterial artificial chromosome (BAC). Subsequent screening of the library for key methylotrophy genes resulted in the discovery of a clone carrying a *pmoCAB* operon, encoding for the particulate methane monooxygenase. A complete sequence of the operon was determined by shotgun sequencing. The sequence of the identified *pmoA* gene was almost identical to a *Methylocystis* sp. sequence which had been previously detected in this soil (Radajewski et al., 2002). Additionally, 12 other putative genes were detected on the same clone (Dumont et al., 2006), 3 of which take part in C_1 metabolism. Labeled methylotroph populations were also analyzed phylogenetically by 16S rRNA gene DGGE fingerprints and subsequent sequencing of dominant DGGE bands. In addition to *Methylocystis*, the methanotrophic genera *Methylobacter* and *Methylocella* were identified together with sequences similar to *Bacteroidetes* and *γ-Proteobacteria*.

One of the main criticisms associated with SIP is considered to be inappropriately high concentrations of labeled substrates introduced during incubations compared to concentrations that occur *in situ*. However, these high concentrations were necessary to achieve sufficient yields of labeled DNA for metagenomic analysis. Murrell and colleagues were the first to resolve this issue (Neufeld et al., 2008). Their strategy was the application of multiple displacement amplification to "bridge the gap" between the picogram quantities of labeled DNA and required microgram quantities for subsequent metagenomic analyses. Their study found *Methylophaga* spp. to be involved in oceanic methanol cycling and detected a 9-kb DNA fragment that encoded for the enzymes involved in methanol dehydrogenase synthesis, regulation, and assembly. Similarly, the techniques described were used for the analysis of as yet uncultured *Methylocystis*-related populations in acidic peatlands (Chen et al., 2008). These populations, which had been found to be dominant in the majority of acidic peatlands sampled, were further confirmed to be actively involved in oxidizing methane by SIP-based investigations. After $^{13}CH_4$ had been assimilated, ^{13}C-labeled DNA containing the

genome of *Methylocystis* spp. was used for a construction of a metagenomic library and screened for the presence of key methylotrophic genes. Shotgun sequencing of a clone containing methanol dehydrogenase gene permitted the researchers to assemble a gene cluster encoding polypeptides involved in methanol utilization (*mxaFJGIRSAC*). These reports (Dumont et al., 2006, Chen et al., 2008, Neufeld et al., 2008) were the first ones to show that retrieval of targeted genetic information can be achieved with minimal sequencing effort. At the same time, the authors proposed early sequencing of complete genomes of microbial populations directly from the environment (Dumont et al., 2006). Not long after, a nearly complete genome was obtained of a novel uncultured methylotrophic bacterium *Methylotenera mobilis* from the water and sediments of Lake Washington, WA, USA (Kalyuzhnaya et al., 2008). This proof-of-principle study shows that genomes of ecologically relevant subpopulations can be reassembled after whole genome shotgun sequencing of stable isotopically labeled DNA. Additional results of this study revealed several clades of bacteria involved in C_1 substrate metabolism; some were traditional methylotrophs such as *Methylobacter*, *Methylomonas*, *Methyloversatilis*, or *Ralstonia*, and others clustered with *Verrucomicrobia*, *Nitrospirae*, and *Planctomycetes*, clades not commonly associated with methylotrophs.

SIP and metagenomics to study biodegradation of anthropogenic compounds

Some of the most widespread and environmentally significant xenobiotics are polychlorinated biphenyls (PCBs) (Breivik et al., 2002). Correspondingly, studies combining SIP and metagenomics to study PCB-degrading bacteria are common. The first such study performed DNA-SIP integrated with GeoChip-mediated metagenomic analysis of bacteria in the root zone of an Austrian pine (*Pinus nigra* L.) growing naturally in PCB-contaminated soil using ^{13}C-biphenyl, a PCB analogue, as a substrate (Leigh et al., 2007). The findings of this study pointed to novel populations of biphenyl-utilizing bacteria, including

Pseudonocardia, *Nocardioides*, *Kribella*, *Variovorax*, and *Polaromonas* in addition to previously known PCB-degrading *Sphingomonas* spp. GeoChip analyses of ^{13}C-DNA detected 30 genes associated with organic contaminant degradation in the ^{13}C-DNA, majority of which were associated with the degradation of aromatics, including biphenyl, benzoate, catechol, protocatechuate, naphthalene, phenol, dibenzofuran, and phenylpropionate. The presence of these genes in biphenyl-labeled populations suggests that they have the potential to degrade several aromatic substrates. In addition, genes of the β-ketoadipate pathway were detected indicating potential abilities of the populations to mineralize monoaromatics once biphenyl has been transformed into monoaromatic intermediates. Only four genes, however, were detected from the biphenyl upper pathway (*bph* operon) associated with rhodococci and bacilli, suggesting that only a tiny fraction of the actual diversity in upper biphenyl pathway genes had been revealed. This hypothesis was supported by PCR amplification and sequence analyses of genes encoding for aromatic ring hydroxylating dioxygenases (ARHD), all of which shared homology but were not identical to those previously deposited in GenBank. Thus, these novel genes were undetected using the microarray. Moreover, some of the sequences did not cluster with any known ARHDs and represented a novel clade.

Sul and colleagues (2009) applied metagenomics directly to isolate a novel biphenyl dioxygenase (*bphA*) gene from PCB-contaminated river sediment bacteria enriched by the incubation with ^{13}C-biphenyl. Biphenyl dioxygenase (BphAE), a multicomponent enzyme catalyzing the activation of biphenyl ring by insertion of two oxygen atoms, is crucial for biodegradation of biphenyl. Degradation of PCBs is permitted by relaxed substrate specifity of the enzyme, which has been determined to be closely connected with its primary structure (Vézina et al., 2007, 2008). The dioxygenase sequence detected by Sul et al. (2009) was highly similar to that in *Pseudomonas* sp. Cam-1 and *Pseudomonas pseudoalcaligenes* KF707. Although in most laboratory PCB-degrading strains, genes *bphAE* are organized in

operons with other enzymes for subsequent transformation of dixydroxylated biphenyl, this clone only contained *bphAE* genes. The authors ascribe this phenomenon to an acquisition of the genes from another microorganism during exposure of the sediment to PCBs, possibly by horizontal gene transfer. This hypothesis was supported by different G+C content of *bphAE* than average for the cloned fragment. The activity of BphAE was tested after expression of the genes *bphAE* from the cosmid clone along with *bphFGBC* from *Burkholderia xenovorans* LB400. The spectrum of metabolized PCB congeners was similar to that of *P. pseudoalcaligenes* KF707, transforming only the congeners without chloro substitutions at the 2,3 positions.

The identification of previously characterized organisms and genes by Sul et al. (2009) may be due to the long SIP incubation and enrichment of fast-growing organisms that would be amenable to detection using cultivation-based approaches. In the future, combining direct shotgun metagenomic sequencing or function-based screening with DNA-SIP could enable the discovery of highly novel degradative genes, rather than those with similarity to known sequences as were detected with Leigh et al. (2007) and Sul et al. (2009), which is necessary for PCR-based or microarray-based detection. Although Sul et al. (2009) did not discover of a dioxygenase with a broader substrate specificity than had been observed previously, this study contributes to our understanding of genomic features of degradative populations. This particular case suggests that the gene organization of *bph* genes in nature might be scattered rather than clustered in operons. The idea of catabolic genes being dispersed on chromosomes and plasmids was supported by another paper which discusses the organization of aromatic degradation pathway genes (Suenaga et al., 2009). Thirty-eight fosmid clones were analyzed carrying genes for extradiol dioxygenases, and only two of the clones contained complete degradation pathways that are commonly found in aromatic compound-utilizing isolates. The other clones contained only subsets of the pathway genes with novel gene arrangements.

Recent results (Uhlík et al., 2012) demonstrated the ability of biphenyl-metabolizing bacteria to utilize other aromatic compounds in contaminated soil; populations of *Rhodanobacter*, *Burkholderia*, *Pandoraea*, *Dyella* and other *Proteobacteria* were observed to derive carbon from benzoate and naphthalene in addition to biphenyl. This study combined SIP with sequence analysis of 16S rRNA gene pyrotags amplified from ^{13}C-DNA to identify taxa associated with the biodegradation of pollutants. Results of a few recently published bioremediation-related SIP studies reveal bacteria that had not been associated with utilization of the substrates before. Examples include newly associated populations of *Pusillimonas* or *Rhodanobacter* with the degradation of biphenyl (Lee et al., 2011, Uhlík et al., 2012) and *Thermincola* with the degradation of toluene (Pilloni et al., 2011). In addition, many unclassified 16S rRNA gene sequences were retrieved from ^{13}C-DNA labeled by different substrates pointing to novel yet-to-be described bacterial taxa involved in biodegradation of biphenyl, benzoate, naphthalene, or toluene (Lee et al., 2011, Pilloni et al., 2011, Uhlík et al., 2012).

POTENTIAL CONTRIBUTIONS OF SIP AND METAGENOMICS TO BIOREMEDIATION

The fundamental research on microbial aspects of bioremediation improves understandings of processes and could potentially improve bioremediation technologies. Thorough analysis of the labeled metagenomes of bioremediative populations can provide valuable information for assessing bioremediation potential of autochthonous microorganisms as well as designing and monitoring engineered bioremediation strategies.

Assessing bioremediation potential

Before bioremediation strategies are applied to a contaminated site, the bioremediation potential of the indigenous microflora should be assessed. In this case, SIP can be valuable

for determining whether organisms capable of metabolizing the contaminant are already present at the site. If so, then biostimulation would likely be a viable bioremediation strategy. For these purposes, SIP incubations can be performed either *in vitro* using microcosms constructed from field-collected samples (Leigh et al., 2007, Uhlík et al., 2009b, Winderl et al., 2010) or directly *in situ* (Padmanabhan et al., 2003, DeRito et al., 2005, Mahmood et al., 2005, Liou et al., 2008, Pumphrey and Madsen, 2008, Bombach et al., 2010). Metagenomic functional gene analyses of SIP studies is particularly valuable in the case of biodegradation of xenobiotics that occur as mixtures, such as PCBs, since biphenyl dioxygenase enzymes vary widely in their capability to degrade different lower chlorinated PCB congeners (Erickson and Mondello, 1993, Mondello et al., 1997). Analyzing the SIP-labeled dioxygenase gene sequence in relation to known enzymes with known substrate specificities can help predict which of the congeners are likely to be degraded by the microbial community (Barriault et al., 2002, Vézina et al., 2008). The absence or inactivity of dioxygenases with appropriate congener specificity for the PCBs on-site might indicate that an anaerobic treatment to promote dehalogenation would be appropriate (Wiegel and Wu, 2000, Smidt and de Vos, 2004). In addition, metagenomic exploration of active populations can clarify the metabolic capabilities as well as regulatory mechanisms within microbes, such as through sequence-based detection of regulatory elements, and these properties can be subjected to genetic manipulations with the aim of improving the efficacy of bioremediation through bioaugmentation.

Bioaugmentation

SIP is ideal for identifying microbes optimal for bioaugmentation for several reasons. First, SIP-metagenomics has the potential to reveal the identity of microbial species that are metabolically active under a variety of environmental conditions (type of matrix, temperature, moisture, oxygen levels, etc.) or that derive carbon from more than one

contaminant. Second, once bioaugmentation with exogenous organisms has been applied on a contaminated site, SIP can help determine whether the added microorganisms have adapted to new conditions and survived and, more importantly, if they are actively biodegrading the contaminant(s).

Understanding the mechanisms underlying biostimulation and phytoremediation technologies

Biostimulation methods seek to address limitations in environmental conditions and/or bioavailability in order to enable indigenous microbial communities to more rapidly biodegrade the contaminant. One approach to biostimulation is the use of plants to promote biodegradative activity in the root zone, which is a form of phytoremediation known as rhizoremediation (Kuiper et al., 2004, Macková et al., 2006, Gerhardt et al., 2009, Macek et al., 2009). Plants can rhizostimulate contaminant biodegradation through a variety of mechanisms. Root exudation and root decay provides rhizosphere microorganisms with growth substrates and secondary compounds that may function as cometabolites or inducers of biodegradative pathways, or surfactants and phytochemicals that increase the bioavailability of poorly soluble pollutants (Leigh et al., 2002, Bertin et al., 2003, Singer et al., 2003, Singer et al., 2004, Leigh et al., 2006, Yi and Crowley, 2007, Toussaint et al., 2012). In addition, the rhizosphere is also richer in oxygen essential for the activity of aerobic organisms and their dioxygenases and monooxygenases that are often involved in biodegradation processes of organic contaminants (Leigh et al., 2002). Understanding how the microbial community responds to biostimulation is important to improving bioremediation technologies. SIP and metagenomics approaches could provide novel and important insight into the mechanisms underlying rhizoremediation (Prosser et al., 2006) by identifying contaminant degraders active under different environmental and biostimulatory conditions.

Plants also host many bacteria in their endosphere. Emerging research indicates that they contribute to biodegradation of toxic organic compounds in contaminated soil and could have potential for improving phytoremediation (Newman and Reynolds, 2005, McGuinness and Dowling, 2009, Weyens et al., 2009, Weyens et al., 2011). Although no studies have as of yet employed SIP-metagenomics to study microbial metabolism inside plants, this topic is likely to be investigated with respect to bioremediation in the future.

Some plants contribute to remediation more directly, by taking up and transforming contaminants or storing them in above-ground parts (Macek et al., 2000). Studies have been published showing that plants, however, have limited abilities to mineralize pollutants. In case of PCBs, for instance, monohydroxylated and/or dihydroxylated derivatives are formed without the aromatic ring being cleaved (Rezek et al., 2008). In connection with improving phytoremediation efficiency, novel genes revealed by SIP-metagenomics can be used in order to prepare genetically modified plants with abilities to cleave aromatic structures (Macek et al., 2008, Nováková et al., 2009, Sylvestre et al., 2009, Van Aken et al., 2010).

SIP appears to be an ideal tool for use during laboratory tests evaluating how different biostimulation strategies alter biodegradative microbial populations, and the resultant information could also potentially be useful for creating and monitoring engineered bioremediation treatments.

Carbon flow through contaminated systems

One of the major drawbacks of DNA-SIP is the detection of cross-feeding populations in addition to primary utilizers of a substrate (Neufeld et al., 2007c). Experimental data have shown a very rapid sequestration of carbon in some microbial communities (Lueders et al., 2006). In addition, mineralization of ^{13}C-labeled substrates generates $^{13}CO_2$, which can further be taken up by autotrophic bacteria and incorporated into their biomass biasing thereby the conclusions of a study. Therefore, different strategies have been suggested to

minimize the effect of cross-feeding detection by DNA-SIP or to clearly distinguish it from primary utilization. These include time-course experiments (Leigh et al., 2007), shortening the incubation times and addition of ^{13}C-carrier DNA in the density gradient separation (Gallagher et al., 2005), or combination of DNA-SIP with RNA-SIP (Dumont et al., 2011). Yet in the context of bioremediation, cross-feeding populations may also be important contributors. For instance, intermediate metabolites or incompletely-degraded substrates released by the biodegradative population may be removed from the environment by cross-feeding microbes. Monitoring cross-feeding also has the potential to identify all microorganisms essential to complete degradation of contaminants and to reveal the flow of carbon through the microbial food web in contaminated environments.

Recently, a new SIP methodology has been developed which is ideal for determining carbon flow through food webs. Known as stable isotope switching (SIS) (Maxfield et al., 2012), the approach is based on the incubation of the matrix with ^{13}C-labeled substrate that is switched for unlabeled substrate when full labeling is achieved. Such a setting allows for monitoring carbon uptake, turnover, release, and sequestration.

LIMITATIONS

Despite the great potential contribution of SIP and metagenomics to microbial ecology and biotechnology, one has to realistically evaluate the limitations associated with each of the approaches. A major challenge associated with SIP is a very limited availability and high cost of labeled substrates. SIP is also very labor-intensive and low-throughput using current techniques. These factors may prohibit performing a sufficient number of replicates to enable statistically valid comparisons among treatments. In the future, the use of robotics and automation could assist in increasing the throughput of SIP gradient fractionation and assessments. Another challenge of SIP is the need to combine many techniques and

technologies spanning many disciplines, which may be an obstacle for researchers without experience in these techniques (reviewed by Chen and Murrell, 2010).

When designing SIP-based experiments, it is also important to use realistic substrate concentrations and incubation times that are of appropriate duration to enable labeling yet minimize potential cross-feeding and over-enrichment (Neufeld et al., 2007a). Inappropriately high concentrations of a labeled substrate can lead to population shifts, such as opportunistic growth of some degraders, suppression of populations adapted to only low concentrations of the substrate, or shifts caused by accumulation of inhibitory or toxic intermediates.

It has been proposed earlier (Madsen, 2006) that, especially for bioremediation purposes, SIP assay conditions should ideally match the conditions in the field. Field conditions are constantly changing (temperature, humidity, etc.) and may be affected by larger scale processes such as living plant roots, and these factors are challenging to adequately represent under laboratory microcosms. Microcosms often create other unrealistic conditions associated with depletion of oxygen, increased concentrations of CO_2, unrealistic substrate concentrations, or unrealistic availability of the substrate (Friedrich, 2006). Therefore, laboratory microcosms need to be recognized as model systems, with the caveat that they will not fully represent field conditions. Thorough bioremediation surveys require either performing SIP directly *in situ* (Padmanabhan et al., 2003, DeRito et al., 2005, Mahmood et al., 2005, Liou et al., 2008, Pumphrey and Madsen, 2008, Bombach et al., 2010) or a very careful design and monitoring of microcosm experiments.

One of the major limitation of SIP-metagenomics for bioremediation purposes is the restriction of its utility solely to assimilatory processes (Andreoni and Gianfreda, 2007). Yet many bioremediation processes are dissimilatory, for example, with the contaminant serving as the terminal electron acceptor. Many anthropogenic compounds are halogenated

(polychlorinated dioxins, dibenzofurans, biphenyls, or chlorinated solvents), and their bioremediation relies on dehalorespiration (i.e., utilization of a halogenated compound as a terminal electron acceptor to yield energy) (Futagami et al., 2008). These processes can be monitored through different stable isotope analyses (for review, see Ruess and Chamberlain, 2010, Boecklen et al., 2011, Braeckevelt et al., 2012).

Metagenomics techniques are also not exempt from limitations. Rapidly advancing high-throughput sequencing technologies have substantially reduced the cost of sequencing efforts over traditional Sanger sequencing. Sequence-based screening of metagenomic data, however, relies on databases that are far from being completely or accurately annotated. The incompleteness of genetic databases means that about 30% genes sequenced can remain unassigned (Harrington et al., 2007), which prevents metagenomics from achieving its full potential. Other potential problems are associated with the function-based screening of metagenomic libraries. These include the necessity of analyzing a large number of clones to recover positive ones (Sul et al., 2009), possible inabilities of the host cells (usually *E. coli*) to express foreign genes and to form active proteins, or problems associated with cloning of large gene clusters which requires high-molecular weight DNA of high purity (Daniel, 2005, Simon and Daniel, 2009). At the same time, problems associated with gene-targeted metagenomics have been identified. These include mainly errors introduced during PCR by DNA polymerase, generation of chimeric sequences when heterologous templates are used, or errors introduced during pyrosequencing. As a result of these phenomena, diversity is overestimated (Kunin et al., 2010) unless appropriate tools are used for sequence processing. Introduction of several denoising algorithms (Quince et al., 2009, Reeder and Knight, 2010, Quince et al., 2011), improved operational taxonomic unit (OTU) clustering (Huse et al., 2010), or their combination (Schloss et al., 2011, Uhlík et al., 2012) have resulted in much more accurate diversity estimations. Additionally, the most commonly used phylogenetic

marker gene is 16S rRNA, which is present in different copy numbers in different species, ranging from 1 to 15 (Pei et al., 2010). As a result, the relative abundance of 16S rRNA genes in a sequence library does not necessarily correspond to number of cells, and may distort the apparent community structure.

NEW FRONTIERS OF SIP

Realizing the limitations of DNA-SIP and metagenomics gives rise to alternative approaches being considered. One of these is SIP-metatranscriptomics. Using mRNA-SIP benefits from high sensitivity and the fact that mRNA bears functional information. Studies carried out so far by Huang et al. (2009) and Dumont et al. (2011) focused on one transcript each – naphthalene dioxygenase and methane monooxygenase, respectively. The rapidly advancing field of metatranscriptomics (Filiatrault, 2011, Simon and Daniel, 2011, Su et al., 2012) seems to have overcome some of the major obstacles linked with mRNA, such as low recovery of high-quality environmental transcripts, instability of mRNA, or difficulties in separation of mRNA from other RNA species (Mou et al., 2011, Ottesen et al., 2011, Feike et al., 2012, Marchetti et al., 2012, Rinta-Kanto et al., 2012, Stewart et al., 2012). Therefore, it is only a matter of time before mRNA-SIP is combined with metatranscriptomic analyses to provide comprehensive information on the actively transcribed genes associated with the utilization of a specific substrate.

Another group of biomarkers bearing both phylogenetic and functional information are proteins. Protein-SIP was initially applied to track carbon flow in pure cultures (Jehmlich et al., 2008a, Jehmlich et al., 2008b), however the authors proposed its use for the analysis of phylogenetically diverse microbial communities as well, which was reported shortly thereafter (Bastida et al., 2010, Bastida et al., 2011, Pan et al., 2011). More recently, protein-SIP was used for quantitative analysis of induced proteins in substrate shift experiments

(Taubert et al., 2011). Protein-SIP is currently very challenging due to the inability to assign even a putative function to many proteins. Yet protein-SIP is very promising as it may provide a more substantial access to real microbial activity, since proteins are the most explicit indicators of metabolic activity. Therefore, in the future, with more complete databases, better modelling systems, and further development of metaproteomics techniques, protein-SIP is expected to widely expand and advance the field of microbial ecology.

ALTERNATIVES TO SIP

Although SIP is the main focus of this review, other technologies expanding the scope of isotope labeling experiments have been developed that can potentially link microbial phylogeny with metabolic activity (Gutierrez-Zamora and Manefield, 2010). Most of these methods involve fluorescence *in situ* hybridization (FISH), which is a technique commonly applied to taxonomically identify microbial cells using rRNA-targeted oligonucleotide probes (Wagner et al., 2003). FISH developed into a function-identity method after it was combined with (i) microautoradiography (FISH-MAR), which uses radioactive isotopes to monitor microbial uptake of labeled substrates (Lee et al., 1999); (ii) secondary ion mass spectrometry (FISH-SIMS) and nanometer-scale secondary ion mass spectrometry (FISH-nanoSIMS), which determines the isotopic composition of the FISH-identified targeted cells (Orphan et al., 2001, Kuypers and Jørgensen, 2007, Li et al., 2008); (iii) Raman microspectroscopy (Raman-FISH), which detects vibrational shifts of covalent bonds in molecules of FISH-labeled cells (Huang et al., 2007); or (iv) immunomagnetic cell capture (magneto-FISH), which permits targeted magnetic capture of FISH-labeled microorganisms and cell aggregates (Pernthaler et al., 2008). Behrens et al. (2008) reported a new combination of FISH and nanoSIMS for linking microbial phylogeny to metabolic activity at the single cell level by applying enhanced element labeling of microbial cells by FISH (EL-FISH). Such a

modification allowed them to increase the sensitivity of cell detection and broadened the applicability of the methodology for environmental studies. Some of these techniques, mainly FISH-MAR, have been employed in bioremediation studies (Yang et al., 2003, Hesselsoe et al., 2008).

In addition to FISH, other isotope-based techniques have been used such as isotope arrays (Adamczyk et al., 2003) or shotgun isotope arrays (Tobino et al., 2011), small subunit-isotope ratio mass spectrometry (SSU-IRMS) (MacGregor et al., 2002, MacGregor et al., 2006), or radioactive isotope probing (RIP) (Nikolausz et al., 2007). The advantages and limitations of these methods as well as their comparison to SIP are discussed elsewhere (Wagner et al., 2006, Neufeld et al., 2007c, Gutierrez-Zamora and Manefield, 2010).

Recently some FISH-based methods have been combined with SIP in order to link microbial phylogeny to metabolic activity at the single-cell level. For example, Huang and colleagues combined rRNA and mRNA-SIP with single-cell Raman-FISH (Huang et al., 2009). The main advantage of such a combined approach is that it allows for determination and quantification of *in situ* functions of a microbial community. RNA-SIP has also been combined with magnetic-bead capture hybridization (MacGregor et al., 2006, Miyatake et al., 2009). This approach is applicable for linking phylogeny with metabolic activity at the level of class or family while requiring approximately 10,000 times less stable isotope enrichment of RNA, allowing thus the use of environmentally relevant concentrations of the isotopically labeled target substrate. The principle of RNA-SIP combined with magnetic-bead capture hybridization lies in extracting RNA from the matrix after it has been stable isotope labeled, hybridizing the RNA with probes, capturing those hybrids with beads, and collecting beads with a magnet. Captured rRNA is then released and analyzed by isotope ratio mass spectrometry (IRMS). This approach thus eliminates possible biases associated with the use

of artificially high substrate concentrations during incubations relative to those *in situ*, which is one of the major criticisms of SIP (Dumont and Murrell, 2005, Neufeld et al., 2008).

CONCLUSIONS

Since the first development of SIP more than a decade ago (Boschker et al., 1998, Radajewski et al., 2000), this technique has progressed rapidly and has significantly broadened the field of microbial ecology, first by linking phylogenetic identity with function, including of novel clades with no cultured representatives, and later by linking function with novel functional gene sequences. The recent integration of SIP with metagenomics has enabled a more comprehensive understanding of the functional community dynamics of entire microbial systems. An improved mechanistic understanding of microbial ecological function has the potential to enable new breakthroughs in bioremediation technologies such as biostimulation, bioaugmentation and phytoremediation. The great advantage of SIP is its ability to enable a focused detection and analysis of only the organisms active in the utilization of a specific substrate, either directly or indirectly through the food web. In the context of bioremediation, SIP-metagenomics is invaluable for revealing the identity of contaminant-degraders and their functional genes, and this information may be used to assess their response to biostimulation methods or to identify organisms and genes useful for bioaugmentation. As with any method, SIP has limitations, such as false positives, in which cross-feeding organisms may be mistaken for primary utilizers, and the potential for microbial community change during incubations. However if these factors are taken into account, SIP combined with metagenomics has the potential to provide novel insights to the intricate interactions within microbial ecosystems. Limitations like cross-feeding can also be advantageous, providing valuable new insight into the flow of carbon derived from contaminants through the microbial food web. As cultivation-independent methods like SIP

and metagenomic studies advance, it is important to note that cultivation-based techniques are still a crucial means to verify and investigate the physiology and genetics of individual contaminant-degrading microorganisms, to facilitate bioaugmentation, and to enable improved annotation of metagenomic databases. The integration of SIP with the rapidly advancing field of metagenomics is opening new windows into microbial processes and interactions, which may enable major breakthroughs in the field of bioremediation.

In loving memory of Prof. Martina Macková (May 7, 1965 – August 2, 2012).

ACKNOWLEDGEMENTS

The authors acknowledge the support of Czech Ministry of Education, Youth and Sports (grant ME09024), and 7th Framework Programme (grant MINOTAURUS, no: 265946). This publication was made possible by grants from the national center for research resources (5P20RR016466-12) and the national institute of general medical sciences (8P20GM103395-12) from the national institutes of health. MCL acknowledges the David L. Boren National Security Education Program (NSEP) Fellowship for support.

REFERENCES

Aanderud ZT, Lennon JT. Validation of heavy-water stable isotope probing for the characterization of rapidly responding soil bacteria. Appl Environ Microbiol. 2011;77:4589-96.

Adamczyk J, Hesselsoe M, Iversen N, Horn M, Lehner A, Nielsen PH, Schloter M, Roslev P, Wagner M. The isotope array, a new tool that employs substrate-mediated labeling of rRNA for determination of microbial community structure and function. Appl Environ Microbiol. 2003;69:6875-87.

Andreoni V, Gianfreda L. Bioremediation and monitoring of aromatic-polluted habitats. Appl Microbiol Biotechnol. 2007;76:287-308.

Arai K, Tsubone T, Takechi T, Inoue T. Bioremediation of trichloroethylene and *cis*-1,2-dichloroethylene-contaminated groundwater by methane-utilizing bacteria. J Vet Med Sci. 1999;61:861-3.

Barriault D, Plante MM, Sylvestre M. Family shuffling of a targeted *bphA* region to engineer biphenyl dioxygenase. J Bacteriol. 2002;184:3794-800.

Bastida F, Jechalke S, Bombach P, Franchini AG, Seifert J, von Bergen M, Vogt C, Richnow HH. Assimilation of benzene carbon through multiple trophic levels traced by different stable isotope probing methodologies. FEMS Microbiol Ecol. 2011;77:357-69.

Bastida F, Rosell M, Franchini AG, Seifert J, Finsterbusch S, Jehmlich N, Jechalke S, Von Bergen M, Richnow HH. Elucidating MTBE degradation in a mixed consortium using a multidisciplinary approach. FEMS Microbiol Ecol. 2010;73:370-84.

Behrens S, Losekann T, Pett-Ridge J, Weber PK, Ng WO, Stevenson BS, Hutcheon ID, Relman DA, Spormann AM. Linking microbial phylogeny to metabolic activity at the single-cell level by using enhanced element labeling-catalyzed reporter deposition fluorescence in situ hybridization (EL-FISH) and NanoSIMS. Appl Environ Microbiol. 2008;74:3143-50.

Bell TH, Yergeau E, Martineau C, Juck D, Whyte LG, Greer CW. Identification of nitrogen-incorporating bacteria in petroleum-contaminated Arctic soils by using ^{15}N DNA-based stable isotope probing and pyrosequencing. Appl Environ Microbiol. 2011;77:4163-71.

Bertin C, Yang XH, Weston LA. The role of root exudates and allelochemicals in the rhizosphere. Plant Soil. 2003;256:67-83.

Binga EK, Lasken RS, Neufeld JD. Something from (almost) nothing: the impact of multiple displacement amplification on microbial ecology. ISME J. 2008;2:233-41.

Boecklen WJ, Yarnes CT, Cook BA, James AC. On the use of stable isotopes in trophic ecology. Annu Rev Ecol Evol Syst. 2011;42:411-40.

Bombach P, Chatzinotas A, Neu TR, Kästner M, Lueders T, Vogt C. Enrichment and characterization of a sulfate-reducing toluene-degrading microbial consortium by combining *in situ* microcosms and stable isotope probing techniques. FEMS Microbiol Ecol. 2010;71:237-46.

Boschker HTS, Nold SC, Wellsbury P, Bos D, de Graaf W, Pel R, Parkes RJ, Cappenberg TE. Direct linking of microbial populations to specific biogeochemical processes by ^{13}C-labelling of biomarkers. Nature. 1998;392:801-5.

Braeckevelt M, Fischer A, Kaestner M. Field applicability of Compound-Specific Isotope Analysis (CSIA) for characterization and quantification of in situ contaminant degradation in aquifers. Appl Microbiol Biotechnol. 2012;94:1401-21.

Breivik K, Sweetman A, Pacyna JM, Jones KC. Towards a global historical emission inventory for selected PCB congeners - a mass balance approach. 1. Global production and consumption. Sci Total Environ. 2002;290:181-98.

Brodie EL, Desantis TZ, Joyner DC, Baek SM, Larsen JT, Andersen GL, Hazen TC, Richardson PM, Herman DJ, Tokunaga TK, Wan JM, Firestone MK. Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. Appl Environ Microbiol. 2006;72:6288-98.

Buckley DH, Huangyutitham V, Hsu SF, Nelson TA. Stable isotope probing with $^{15}N_2$ reveals novel noncultivated diazotrophs in soil. Appl Environ Microbiol. 2007a;73:3196-204.

Buckley DH, Huangyutitham V, Hsu SF, Nelson TA. Stable isotope probing with ^{15}N achieved by disentangling the effects of genome G+C content and isotope enrichment on DNA density. Appl Environ Microbiol. 2007b;73:3189-95.

Chen Y, Dumont MG, Neufeld JD, Bodrossy L, Stralis-Pavese N, McNamara NP, Ostle N, Briones MJ, Murrell JC. Revealing the uncultivated majority: combining DNA stable-isotope probing, multiple displacement amplification and metagenomic analyses of uncultivated *Methylocystis* in acidic peatlands. Environ Microbiol. 2008;10:2609-22.

Chen Y, Murrell JC. When metagenomics meets stable-isotope probing: progress and perspectives. Trends Microbiol. 2010;18:157-63.

Daniel R. The metagenomics of soil. Nat Rev Microbiol. 2005;3:470-8.

DeAngelis KM, Wu CH, Beller HR, Brodie EL, Chakraborty R, DeSantis TZ, Fortney JL, Hazen TC, Osman SR, Singer ME, Tom LM, Andersen GL. PCR Amplification-Independent Methods for Detection of Microbial Communities by the High-Density Microarray PhyloChip. Appl Environ Microbiol. 2011;77:6313-22.

DeRito CM, Pumphrey GM, Madsen EL. Use of field-based stable isotope probing to identify adapted populations and track carbon flow through a phenol-degrading soil microbial community. Appl Environ Microbiol. 2005;71:7858-65.

Diaz E. Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility. Int Microbiol. 2004;7:173-80.

Dumont MG, Murrell JC. Stable isotope probing - linking microbial identity to function. Nat Rev Microbiol. 2005;3:499-504.

Dumont MG, Pommerenke B, Casper P, Conrad R. DNA-, rRNA- and mRNA-based stable isotope probing of aerobic methanotrophs in lake sediment. Environ Microbiol. 2011;13:1153-67.

Dumont MG, Radajewski SM, Miguez CB, McDonald IR, Murrell JC. Identification of a complete methane monooxygenase operon from soil by combining stable isotope probing and metagenomic analysis. Environ Microbiol. 2006;8:1240-50.

Dunford EA, Neufeld JD. DNA stable-isotope probing (DNA-SIP). J Vis Exp. 2010.

Edwards RA, Rodriguez-Brito B, Wegley L, Haynes M, Breitbart M, Peterson DM, Saar MO, Alexander S, Alexander EC, Rohwer F. Using pyrosequencing to shed light on deep mine microbial ecology. BMC Genomics. 2006;7.

Erickson BD, Mondello FJ. Enhanced biodegradation of polychlorinated biphenyls after site-directed mutagenesis of a biphenyl dioxygenase gene. Appl Environ Microbiol. 1993;59:3858-62.

Feike J, Jurgens K, Hollibaugh JT, Kruger S, Jost G, Labrenz M. Measuring unbiased metatranscriptomics in suboxic waters of the central Baltic Sea using a new in situ fixation system. ISME J. 2012;6:461-70.

Filiatrault MJ. Progress in prokaryotic transcriptomics. Curr Opin Microbiol. 2011;14:579-86.

Friedrich MW. Stable-isotope probing of DNA: insights into the function of uncultivated microorganisms from isotopically labeled metagenomes. Curr Opin Biotechnol. 2006;17:59-66.

Futagami T, Goto M, Furukawa K. Biochemical and genetic bases of dehalorespiration. Chem Rec. 2008;8:1-12.

Gallagher E, McGuinness L, Phelps C, Young LY, Kerkhof LJ. ^{13}C-carrier DNA shortens the incubation time needed to detect benzoate-utilizing denitrifying bacteria by stable-isotope probing. Appl Environ Microbiol. 2005;71:5192-6.

Geoffrey MG. Metals, minerals and microbes: geomicrobiology and bioremediation. Microbiology. 2010;156:609-43.

Gerhardt KE, Huang XD, Glick BR, Greenberg BM. Phytoremediation and rhizoremediation of organic soil contaminants: Potential and challenges. Plant Sci. 2009;176:20-30.

Gutierrez-Zamora ML, Manefield M. An appraisal of methods for linking environmental processes to specific microbial taxa. Rev Environ Sci Biotechnol. 2010;9:153-85.

Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol. 1998;5:R245-R9.

Harrington ED, Singh AH, Doerks T, Letunic I, von Mering C, Jensen LJ, Raes J, Bork P. Quantitative assessment of protein function prediction from metagenomics shotgun sequences. Proc Natl Acad Sci USA. 2007;104:13913-8.

Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM, Singh N, Jansson JK, Probst A, Borglin SE, Fortney JL, Stringfellow WT, Bill M, Conrad ME, Tom LM, Chavarria KL, Alusi TR, Lamendella R, Joyner DC, Spier C, Baelum J, Auer M, Zemla ML, Chakraborty R, Sonnenthal EL, D'Haeseleer P, Holman HYN, Osman S, Lu ZM, Van Nostrand JD, Deng Y, Zhou JZ, Mason OU. Deep-sea oil plume enriches indigenous oil-degrading bacteria. Science. 2010;330:204-8.

He Z, Deng Y, Van Nostrand JD, Tu Q, Xu M, Hemme CL, Li X, Wu L, Gentry TJ, Yin Y, Liebich J, Hazen TC, Zhou J. GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. ISME J. 2010;4:1167-79.

He Z, Gentry TJ, Schadt CW, Wu L, Liebich J, Chong SC, Huang Z, Wu W, Gu B, Jardine P, Criddle C, Zhou J. GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J. 2007;1:67-77.

Hesselsoe M, Bjerring ML, Henriksen K, Loll P, Nielsen JL. Method for measuring substrate preferences by individual members of microbial consortia proposed for bioaugmentation. Biodegradation. 2008;19:621-33.

Huang WE, Ferguson A, Singer AC, Lawson K, Thompson IP, Kalin RM, Larkin MJ, Bailey MJ, Whiteley AS. Resolving genetic functions within microbial populations: in situ analyses using rRNA and mRNA stable isotope probing coupled with single-cell raman-fluorescence in situ hybridization. Appl Environ Microbiol. 2009;75:234-41.

Huang WE, Stoecker K, Griffiths R, Newbold L, Daims H, Whiteley AS, Wagner M. Raman-FISH: combining stable-isotope Raman spectroscopy and fluorescence *in situ* hybridization for the single cell analysis of identity and function. Environ Microbiol. 2007;9:1878-89.

Huse SM, Welch DM, Morrison HG, Sogin ML. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environ Microbiol. 2010;12:1889-98.

Iwai S, Chai B, Jesus EdC, Penton CR, Lee TK, Cole JR, Tiedje JM. Gene-Targeted Metagenomics (GT Metagenomics) to Explore the Extensive Diversity of Genes of Interest in Microbial Communities. Handbook of Molecular Microbial Ecology I: John Wiley & Sons, Inc.; 2011. p. 235-43.

Jehmlich N, Schmidt F, Hartwich M, von BM, Richnow HH, Vogt C. Incorporation of carbon and nitrogen atoms into proteins measured by protein-based stable isotope probing (Protein-SIP). Rapid Commun Mass Spectrom. 2008a;22:2889-97.

Jehmlich N, Schmidt F, Taubert M, Seifert J, von BM, Richnow HH, Vogt C. Comparison of methods for simultaneous identification of bacterial species and determination of metabolic activity by protein-based stable isotope probing (Protein-SIP) experiments. Rapid Commun Mass Spectrom. 2009;23:1871-8.

Jehmlich N, Schmidt F, von BM, Richnow HH, Vogt C. Protein-based stable isotope probing (Protein-SIP) reveals active species within anoxic mixed cultures. ISME J. 2008b;2:1122-33.

Jørgensen KS. *In Situ* Bioremediation. Adv Appl Microbiol. 2007;61:285-305.

Kalyuzhnaya MG, Lapidus A, Ivanova N, Copeland AC, McHardy AC, Szeto E, Salamov A, Grigoriev IV, Suciu D, Levine SR, Markowitz VM, Rigoutsos I, Tringe SG, Bruce DC, Richardson PM, Lidstrom ME, Chistoserdova L. High-resolution metagenomics targets specific functional types in complex microbial communities. Nat Biotechnol. 2008;26:1029-34.

Kane SR, Chakicherla AY, Chain PS, Schmidt R, Shin MW, Legler TC, Scow KM, Larimer FW, Lucas SM, Richardson PM, Hristova KR. Whole-genome analysis of the methyl *tert*-butyl ether-degrading beta-proteobacterium *Methylibium petroleiphilum* PM1. J Bacteriol. 2007;189:1931-45.

Kuiper I, Lagendijk EL, Bloemberg GV, Lugtenberg BJ. Rhizoremediation: a beneficial plant-microbe interaction. Mol Plant Microbe Interact. 2004;17:6-15.

Kunin V, Engelbrektson A, Ochman H, Hugenholtz P. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. Environ Microbiol. 2010;12:118-23.

Kuypers MMM, Jørgensen BB. The future of single-cell environmental microbiology. Environ Microbiol. 2007;9:6-7.

Lee N, Nielsen PH, Andreasen KH, Juretschko S, Nielsen JL, Schleifer KH, Wagner M. Combination of fluorescent in situ hybridization and microautoradiography - a new tool for structure-function analyses in microbial ecology. Appl Environ Microbiol. 1999;65:1289-97.

Lee TK, Lee J, Sul WJ, Iwai S, Chai BC, Tiedje JM, Park J. Novel biphenyl-oxidizing bacteria and dioxygenase genes from a Korean tidal mudflat. Appl Environ Microbiol. 2011;77:3888-91.

Leigh MB, Fletcher JS, Fu X, Schmitz FJ. Root turnover: an important source of microbial substrates in rhizosphere remediation of recalcitrant contaminants. Environ Sci Technol. 2002;36:1579-83.

Leigh MB, Pellizari VH, Uhlík O, Sutka R, Rodrigues J, Ostrom NE, Zhou J, Tiedje JM. Biphenyl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). ISME J. 2007;1:134-48.

Leigh MB, Prouzová P, Macková M, Macek T, Nagle DP, Fletcher JS. Polychlorinated biphenyl (PCB)-degrading bacteria associated with trees in a PCB-contaminated site. Appl Environ Microbiol. 2006;72:2331-42.

Li T, Wu TD, Mazeas L, Toffin L, Guerquin-Kern JL, Leblon G, Bouchez T. Simultaneous analysis of microbial identity and function using NanoSIMS. Environ Microbiol. 2008;10:580-8.

Liou JS, DeRito CM, Madsen EL. Field-based and laboratory stable isotope probing surveys of the identities of both aerobic and anaerobic benzene-metabolizing microorganisms in freshwater sediment. Environ Microbiol. 2008;10:1964-77.

Little CD, Palumbo AV, Herbes SE, Lidstrom ME, Tyndall RL, Gilmer PJ. Trichloroethylene biodegradation by a methane-oxidizing bacterium. Appl Environ Microbiol. 1988;54:951-6.

Liu S, Suflita JM. Ecology and evolution of microbial populations for bioremediation. Trends Biotechnol. 1993;11:344-52.

Lovley DR. Cleaning up with genomics: applying molecular biology to bioremediation. Nat Rev Microbiol. 2003;1:35-44.

Lozupone CA, Knight R. Species divergence and the measurement of microbial diversity. FEMS Microbiol Rev. 2008;32:557-78.

Lu ZM, Deng Y, Van Nostrand JD, He ZL, Voordeckers J, Zhou AF, Lee YJ, Mason OU, Dubinsky EA, Chavarria KL, Tom LM, Fortney JL, Lamendella R, Jansson JK, D'Haeseleer P, Hazen TC, Zhou JZ. Microbial gene functions enriched in the Deepwater Horizon deep-sea oil plume. ISME J. 2012;6:451-60.

Lueders T, Kindler R, Miltner A, Friedrich MW, Kaestner M. Identification of bacterial micropredators distinctively active in a soil microbial food web. Appl Environ Microbiol. 2006;72:5342-8.

Macek T, Kotrba P, Svatoš A, Nováková M, Demnerová K, Macková M. Novel roles for genetically modified plants in environmental protection. Trends Biotechnol. 2008;26:146-52.

Macek T, Macková M, Káš J. Exploitation of plants for the removal of organics in environmental remediation. Biotechnol Adv. 2000;18:23-34.

Macek T, Uhlík O, Ječná K, Nováková M, Lovecká P, Rezek J, Dudková V, Štursa P, Vrchotová B, Pavlíková D, Demnerová K, Macková M. Advances in phytoremediation and rhizoremediation. In: Varma A, editor. Soil Biology. 17 ed. Berlin, Germany: Springer; 2009. p. 257-77.

MacGregor BJ, Boschker HT, Amann R. Comparison of rRNA and polar-lipid-derived fatty acid biomarkers for assessment of ^{13}C-substrate incorporation by microorganisms in marine sediments. Appl Environ Microbiol. 2006;72:5246-53.

MacGregor BJ, Bruchert V, Fleischer S, Amann R. Isolation of small-subunit rRNA for stable isotopic characterization. Environ Microbiol. 2002;4:451-64.

Macková M, Dowling D, Macek T. Phytoremediation and Rhizoremediation. Theoretical Background. Focus on Biotechnology. Dordrecht, Netherlands: Springer; 2006. p. 300.

Madsen EL. The use of stable isotope probing techniques in bioreactor and field studies on bioremediation. Curr Opin Biotechnol. 2006;17:92-7.

Madsen EL. Stable Isotope Probing Techniques and Bioremediation. In: Murrell JC, Whiteley A, editors. Stable Isotope Probing and Related Technologies. Washington, DC: ASM Press; 2010. p. 165-202.

Mahmood S, Paton GI, Prosser JI. Cultivation-independent in situ molecular analysis of bacteria involved in degradation of pentachlorophenol in soil. Environ Microbiol. 2005;7:1349-60.

Manefield M, Whiteley AS, Bailey MJ. What can stable isotope probing do for bioremediation? Int Biodeterior Biodegrad. 2004;54:163-6.

Manefield M, Whiteley AS, Griffiths RI, Bailey MJ. RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. Appl Environ Microbiol. 2002;68:5367-73.

Marchetti A, Schruth DM, Durkin CA, Parker MS, Kodner RB, Berthiaume CT, Morales R, Allen AE, Armbrust EV. Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. Proc Natl Acad Sci USA. 2012;109:E317-E25.

Maxfield PJ, Dildar N, Hornibrook ERC, Stott AW, Evershed RP. Stable isotope switching (SIS): a new stable isotope probing (SIP) approach to determine carbon flow in the soil food web and dynamics in organic matter pools. Rapid Commun Mass Spectrom. 2012;26:997-1004.

McGuinness M, Dowling D. Plant-associated bacterial degradation of toxic organic compounds in soil. Int J Environ Res Public Health. 2009;6:2226-47.

Miyatake T, MacGregor BJ, Boschker HT. Linking microbial community function to phylogeny of sulfate-reducing *Deltaproteobacteria* in marine sediments by combining stable isotope probing with magnetic-bead capture hybridization of 16S rRNA. Appl Environ Microbiol. 2009;75:4927-35.

Mondello FJ, Turcich MP, Lobos JH, Erickson BD. Identification and modification of biphenyl dioxygenase sequences that determine the specificity of polychlorinated biphenyl degradation. Appl Environ Microbiol. 1997;63:3096-103.

Morales SE, Holben WE. Linking bacterial identities and ecosystem processes: can 'omic' analyses be more than the sum of their parts? FEMS Microbiol Ecol. 2011;75:2-16.

Mou XZ, Vila-Costa M, Sun SL, Zhao WD, Sharma S, Moran MA. Metatranscriptomic signature of exogenous polyamine utilization by coastal bacterioplankton. Environ Microbiol Rep. 2011;3:798-806.

Mulligan CN, Yong RN. Natural attenuation of contaminated soils. Environ Int. 2004;30:587-601.

Nakatsu CH, Hristova K, Hanada S, Meng XY, Hanson JR, Scow KM, Kamagata Y. *Methylibium petroleiphilum* gen. nov., sp. nov., a novel methyl *tert*-butyl ether-degrading methylotroph of the *Betaproteobacteria*. Int J Syst Evol Microbiol. 2006;56:983-9.

Neufeld JD, Chen Y, Dumont MG, Murrell JC. Marine methylotrophs revealed by stable-isotope probing, multiple displacement amplification and metagenomics. Environ Microbiol. 2008;10:1526-35.

Neufeld JD, Dumont MG, Vohra J, Murrell JC. Methodological considerations for the use of stable isotope probing in microbial ecology. Microb Ecol. 2007a;53:435-42.

Neufeld JD, Vohra J, Dumont MG, Lueders T, Manefield M, Friedrich MW, Murrell JC. DNA stable-isotope probing. Nat Protoc. 2007b;2:860-6.

Neufeld JD, Wagner M, Murrell JC. Who eats what, where and when? Isotope-labelling experiments are coming of age. ISME J. 2007c;1:103-10.

Newman LA, Reynolds CM. Bacteria and phytoremediation: new uses for endophytic bacteria in plants. Trends Biotechnol. 2005;23:6-8.

Nikolausz M, Palatinszky MA, Rusznyak A, Richnow HH, Kappelmeyer U, Kastner M. Novel approach using substrate-mediated radiolabelling of RNA to link metabolic function with the structure of microbial communities. FEMS Microbiol Lett. 2007;274:154-61.

Nováková M, Macková M, Chrastilová Z, Viktorová J, Szekeres M, Demnerová K, Macek T. Cloning the bacterial *bphC* gene into *Nicotiana tabacum* to improve the efficiency of PCB phytoremediation. Biotechnol Bioeng. 2009;102:29-37.

Orphan VJ, House CH, Hinrichs KU, McKeegan KD, DeLong EF. Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. Science. 2001;293:484-7.

Ottesen EA, Marin R, Preston CM, Young CR, Ryan JP, Scholin CA, DeLong EF. Metatranscriptomic analysis of autonomously collected and preserved marine bacterioplankton. ISME J. 2011;5:1881-95.

Pace NR, Stahl DA, Lane DJ, Olsen GJ. Analyzing natural microbial populations by rRNA sequences. ASM News. 1985;51:4-12.

Padmanabhan P, Padmanabhan S, DeRito C, Gray A, Gannon D, Snape JR, Tsai CS, Park W, Jeon C, Madsen EL. Respiration of ^{13}C-labeled substrates added to soil in the field and subsequent 16S rRNA gene analysis of ^{13}C-labeled soil DNA. Appl Environ Microbiol. 2003;69:1614-22.

Pan CL, Fischer CR, Hyatt D, Bowen BP, Hettich RL, Banfield JF. Quantitative tracking of isotope flows in proteomes of microbial communities. Mol Cell Proteomics. 2011;10.

Pei AY, Oberdorf WE, Nossa CW, Agarwal A, Chokshi P, Gerz EA, Jin Z, Lee P, Yang L, Poles M, Brown SM, Sotero S, Desantis T, Brodie E, Nelson K, Pei Z. Diversity of 16S rRNA genes within individual prokaryotic genomes. Appl Environ Microbiol. 2010;76:3886-97.

Pernthaler A, Dekas AE, Brown CT, Goffredi SK, Embaye T, Orphan VJ. Diverse syntrophic partnerships from-deep-sea methane vents revealed by direct cell capture and metagenomics. Proc Natl Acad Sci USA. 2008;105:7052-7.

Pilloni G, von Netzer F, Engel M, Lueders T. Electron acceptor-dependent identification of key anaerobic toluene degraders at a tar-oil-contaminated aquifer by Pyro-SIP. FEMS Microbiol Ecol. 2011;78:165-75.

Prosser JI, Rangel-Castro JI, Killham K. Studying plant-microbe interactions using stable isotope technologies. Curr Opin Biotechnol. 2006;17:98-102.

Pumphrey GM, Madsen EL. Field-based stable isotope probing reveals the identities of benzoic acid-metabolizing microorganisms and their in situ growth in agricultural soil. Appl Environ Microbiol. 2008;74:4111-8.

Quince C, Lanzén A, Curtis TP, Davenport RJ, Hall N, Head IM, Read LF, Sloan WT. Accurate determination of microbial diversity from 454 pyrosequencing data. Nat Methods. 2009;6:639-41.

Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ. Removing noise from pyrosequenced amplicons. BMC Bioinformatics. 2011;12.

Radajewski S, Ineson P, Parekh NR, Murrell JC. Stable-isotope probing as a tool in microbial ecology. Nature. 2000;403:646-9.

Radajewski S, McDonald IR, Murrell JC. Stable-isotope probing of nucleic acids: a window to the function of uncultured microorganisms. Curr Opin Biotechnol. 2003;14:296-302.

Radajewski S, Webster G, Reay DS, Morris SA, Ineson P, Nedwell DB, Prosser JI, Murrell JC. Identification of active methylotroph populations in an acidic forest soil by stable-isotope probing. Microbiology. 2002;148:2331-42.

Reeder J, Knight R. Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. Nat Methods. 2010;7:668-9.

Rezek J, Macek T, Macková M, Tříska J, Růžičková K. Hydroxy-PCBs, methoxy-PCBs and hydroxy-methoxy-PCBs: metabolites of polychlorinated biphenyls formed in vitro by tobacco cells. Environ Sci Technol. 2008;42:5746-51.

Riesenfeld CS, Schloss PD, Handelsman J. Metagenomics: genomic analysis of microbial communities. Annu Rev Genet. 2004;38:525-52.

Rinta-Kanto JM, Sun SL, Sharma S, Kiene RP, Moran MA. Bacterial community transcription patterns during a marine phytoplankton bloom. Environ Microbiol. 2012;14:228-39.

Roh H, Yu CP, Fuller ME, Chu KH. Identification of hexahydro-1,3,5-trinitro-1,3,5-triazine-degrading microorganisms via ^{15}N-stable isotope probing. Environ Sci Technol. 2009;43:2505-11.

Ruess L, Chamberlain PM. The fat that matters: Soil food web analysis using fatty acids and their carbon stable isotope signature. Soil Biol Biochem. 2010;42:1898-910.

Schloss PD, Gevers D, Westcott SL. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS ONE. 2011;6:e27310.

Schloss PD, Handelsman J. Biotechnological prospects from metagenomics. Curr Opin Biotechnol. 2003;14:303-10.

Schloss PD, Handelsman J. Metagenomics for studying unculturable microorganisms: cutting the Gordian knot. Genome Biol. 2005;6:229-.

Schmeisser C, Steele H, Streit WR. Metagenomics, biotechnology with non-culturable microbes. Appl Microbiol Biotechnol. 2007;75:955-62.

Schnoor JL, Licht LA, Mccutcheon SC, Wolfe NL, Carreira LH. Phytoremediation of organic and nutrient contaminants. Environ Sci Technol. 1995;29:A318-A23.

Schwarz S, Waschkowitz T, Daniel R. Enhancement of gene detection frequencies by combining DNA-based stable-isotope probing with the construction of metagenomic DNA libraries. World J Microbiol Biotechnol. 2006;22:363-7.

Shigematsu T, Hanada S, Eguchi M, Kamagata Y, Kanagawa T, Kurane R. Soluble methane monooxygenase gene clusters from trichloroethylene-degrading *Methylomonas* sp. strains and detection of methanotrophs during in situ bioremediation. Appl Environ Microbiol. 1999;65:5198-206.

Simon C, Daniel R. Achievements and new knowledge unraveled by metagenomic approaches. Appl Microbiol Biotechnol. 2009;85:265-76.

Simon C, Daniel R. Metagenomic analyses: Past and future trends. Appl Environ Microbiol. 2011;77:1153-61.

Singer AC, Crowley DE, Thompson IP. Secondary plant metabolites in phytoremediation and biotransformation. Trends Biotechnol. 2003;21:123-30.

Singer AC, Thompson IP, Bailey MJ. The tritrophic trinity: a source of pollutant-degrading enzymes and its implications for phytoremediation. Curr Opin Microbiol. 2004;7:239-44.

Smidt H, de Vos WM. Anaerobic microbial dehalogenation. Annu Rev Microbiol. 2004;58:43-73.

Stewart FJ, Ulloa O, DeLong EF. Microbial metatranscriptomics in a permanent marine oxygen minimum zone. Environ Microbiol. 2012;14:23-40.

Su C, Lei LP, Duan YQ, Zhang KQ, Yang JK. Culture-independent methods for studying environmental microorganisms: methods, application, and perspective. Appl Microbiol Biotechnol. 2012;93:993-1003.

Suenaga H, Koyama Y, Miyakoshi M, Miyazaki R, Yano H, Sota M, Ohtsubo Y, Tsuda M, Miyazaki K. Novel organization of aromatic degradation pathway genes in a microbial community as revealed by metagenomic analysis. ISME J. 2009;3:1335-48.

Sul WJ, Park J, Quensen JF, III, Rodrigues JL, Seliger L, Tsoi TV, Zylstra GJ, Tiedje JM. DNA-stable isotope probing integrated with metagenomics for retrieval of biphenyl dioxygenase genes from polychlorinated biphenyl-contaminated river sediment. Appl Environ Microbiol. 2009;75:5501-6.

Sylvestre M, Macek T, Macková M. Transgenic plants to improve rhizoremediation of polychlorinated biphenyls (PCBs). Curr Opin Biotechnol. 2009;20:242-7.

Takeuchi M, Nanba K, Iwamoto H, Nirei H, Kusuda T, Kazaoka O, Owaki M, Furuya K. In situ bioremediation of a *cis*-dichloroethylene-contaminated aquifer utilizing methane-rich groundwater from an uncontaminated aquifer. Water Res. 2005;39:2438-44.

Taubert M, Jehmlich N, Vogt C, Richnow HH, Schmidt F, von Bergen M, Seifert J. Time resolved protein-based stable isotope probing (Protein-SIP) analysis allows quantification of induced proteins in substrate shift experiments. Proteomics. 2011;11:2265-74.

Tobino T, Kurisu F, Kasuga I, Furumai H. Shotgun isotope array for rapid, substrate-specific detection of microorganisms in a microbial community. Appl Environ Microbiol. 2011;77:7430-2.

Topp E, Hanson RS, Ringelberg DB, White DC, Wheatcroft R. Isolation and characterization of an N-methylcarbamate insecticide-degrading methylotrophic bacterium. Appl Environ Microbiol. 1993;59:3339-49.

Torsvik V, Øvreås L. Microbial diversity and function in soil: from genes to ecosystems. Curr Opin Microbiol. 2002;5:240-5.

Toussaint J-P, Pham T, Barriault D, Sylvestre M. Plant exudates promote PCB degradation by a rhodococcal rhizobacteria. Appl Microbiol Biotechnol. 2012; In production. DOI: 10.1007/s00253-011-3824-z.

Uhlík O, Ječná K, Leigh MB, Macková M, Macek T. DNA-based stable isotope probing: a link between community structure and function. Sci Total Environ. 2009a;407:3611-9.

Uhlík O, Ječná K, Macková M, Vlček C, Hroudová M, Demnerová K, Pačes V, Macek T. Biphenyl-metabolizing bacteria in the rhizosphere of horseradish and bulk soil contaminated by polychlorinated biphenyls as revealed by stable isotope probing. Appl Environ Microbiol. 2009b;75:6471-7.

Uhlík O, Wald J, Strejček M, Musilová L, Rídl J, Hroudová M, Vlček Č, Cardenas E, Macková M, Macek T. Identification of bacteria utilizing biphenyl, benzoate, and naphthalene in long-term contaminated soil. PLoS ONE. 2012; In production. DOI: 10.1371/journal.pone.0040653.

Van Aken B, Correa PA, Schnoor JL. Phytoremediation of polychlorinated biphenyls: new trends and promises. Environ Sci Technol. 2010;44:2767-76.

Van Aken B, Yoon JM, Schnoor JL. Biodegradation of nitro-substituted explosives 2,4,6-trinitrotoluene, hexahydro-1,3,5-trinitro-1,3,5-triazine, and octahydro-1,3,5,7-tetranitro-1,3,5-tetrazocine by a phytosymbiotic *Methylobacterium* sp. associated with poplar tissues (*Populus deltoides* x *nigra* DN34). Appl Environ Microbiol. 2004;70:508-17.

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers YH, Smith HO. Environmental genome shotgun sequencing of the Sargasso Sea. Science. 2004;304:66-74.

Vézina J, Barriault D, Sylvestre M. Family shuffling of soil DNA to change the regiospecificity of *Burkholderia xenovorans* LB400 biphenyl dioxygenase. J Bacteriol. 2007;189:779-88.

Vézina J, Barriault D, Sylvestre M. Diversity of the C-terminal portion of the biphenyl dioxygenase large subunit. J Mol Microbiol Biotechnol. 2008;15:139-51.

Vieites JM, Guazzaroni ME, Beloqui A, Golyshin PN, Ferrer M. Metagenomics approaches in systems microbiology. FEMS Microbiol Rev. 2009;33:236-55.

Wagner M, Horn M, Daims H. Fluorescence *in situ* hybridisation for the identification and characterisation of prokaryotes. Curr Opin Microbiol. 2003;6:302-9.

Wagner M, Nielsen PH, Loy A, Nielsen JL, Daims H. Linking microbial community structure with function: fluorescence *in situ* hybridization-microautoradiography and isotope arrays. Curr Opin Biotechnol. 2006;17:83-91.

Warner KL, Larkin MJ, Harper DB, Murrell JC, McDonald IR. Analysis of genes involved in methyl halide degradation in *Aminobacter lissarensis* CC495. FEMS Microbiol Lett. 2005;251:45-51.

Wellington EMH, Berry A, Krsek M. Resolving functional diversity in relation to microbial community structure in soil: exploiting genomics and stable isotope probing. Curr Opin Microbiol. 2003;6:295-301.

Weyens N, Truyens S, Saenen E, Boulet J, Dupae J, Taghavi S, van der Lelie D, Carleer R, Vangronsveld J. Endophytes and their potential to deal with co-contamination of organic

contaminants (toluene) and toxic metals (nickel) during phytoremediation. Int J Phytoremediat. 2011;13:244-55.

Weyens N, van der Lelie D, Taghavi S, Vangronsveld J. Phytoremediation: plant-endophyte partnerships take the challenge. Curr Opin Biotechnol. 2009;20:248-54.

Whiteley AS, Manefield M, Lueders T. Unlocking the 'microbial black box' using RNA-based stable isotope probing technologies. Curr Opin Biotechnol. 2006;17:67-71.

Wiegel J, Wu Q. Microbial reductive dehalogenation of polychlorinated biphenyls. FEMS Microbiol Ecol. 2000;32:1-15.

Winderl C, Penning H, von Netzer F, Meckenstock RU, Lueders T. DNA-SIP identifies sulfate-reducing *Clostridia* as important toluene degraders in tar-oil-contaminated aquifer sediment. ISME J. 2010;4:1314-25.

Woods A, Watwood M, Schwartz E. Identification of a toluene-degrading bacterium from a soil sample through $H_2{}^{18}O$ DNA stable isotope probing. Appl Environ Microbiol. 2011;77:5995-9.

Yang YR, Zarda A, Zeyer J. Combination of microautoradiography and fluorescence in situ hybridization for identification of microorganisms degrading xenobiotic contaminants. Environ Toxicol Chem. 2003;22:2840-4.

Yi H, Crowley DE. Biostimulation of PAH degradation with plants containing high concentrations of linoleic acid. Environ Sci Technol. 2007;41:4382-8.

Yun J, Ryu S. Screening for novel enzymes from metagenome and SIGEX, as a way to improve it. Microb Cell Fact. 2005;4:8.

Zhang L, Xu Z. Assessing bacterial diversity in soil. J Soil Sediment. 2008;8:379-88.

FIGURES

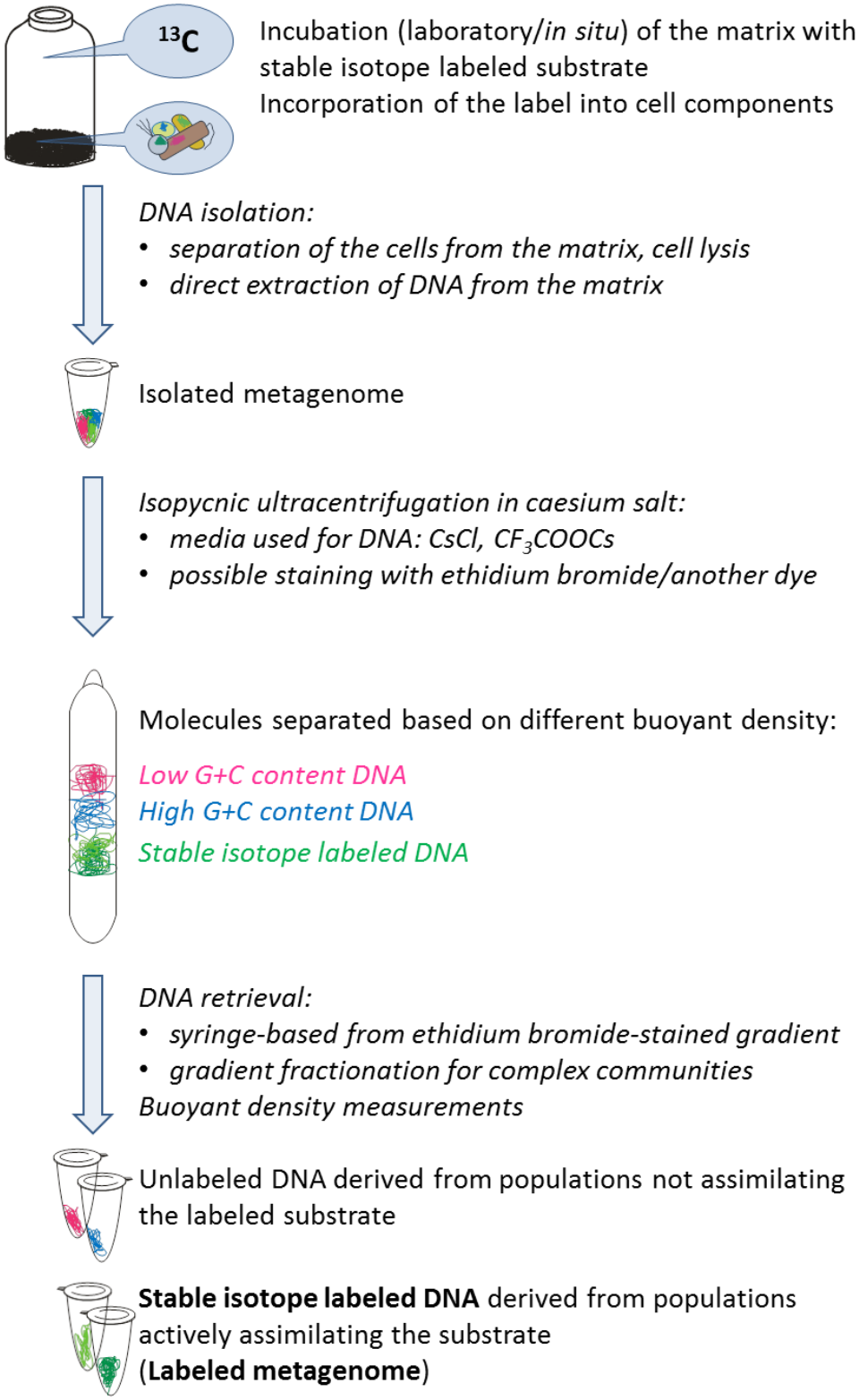

Figure A2-1: Scheme of DNA-based stable isotope probing (SIP) experiment with ^{13}C-labeled substrate.

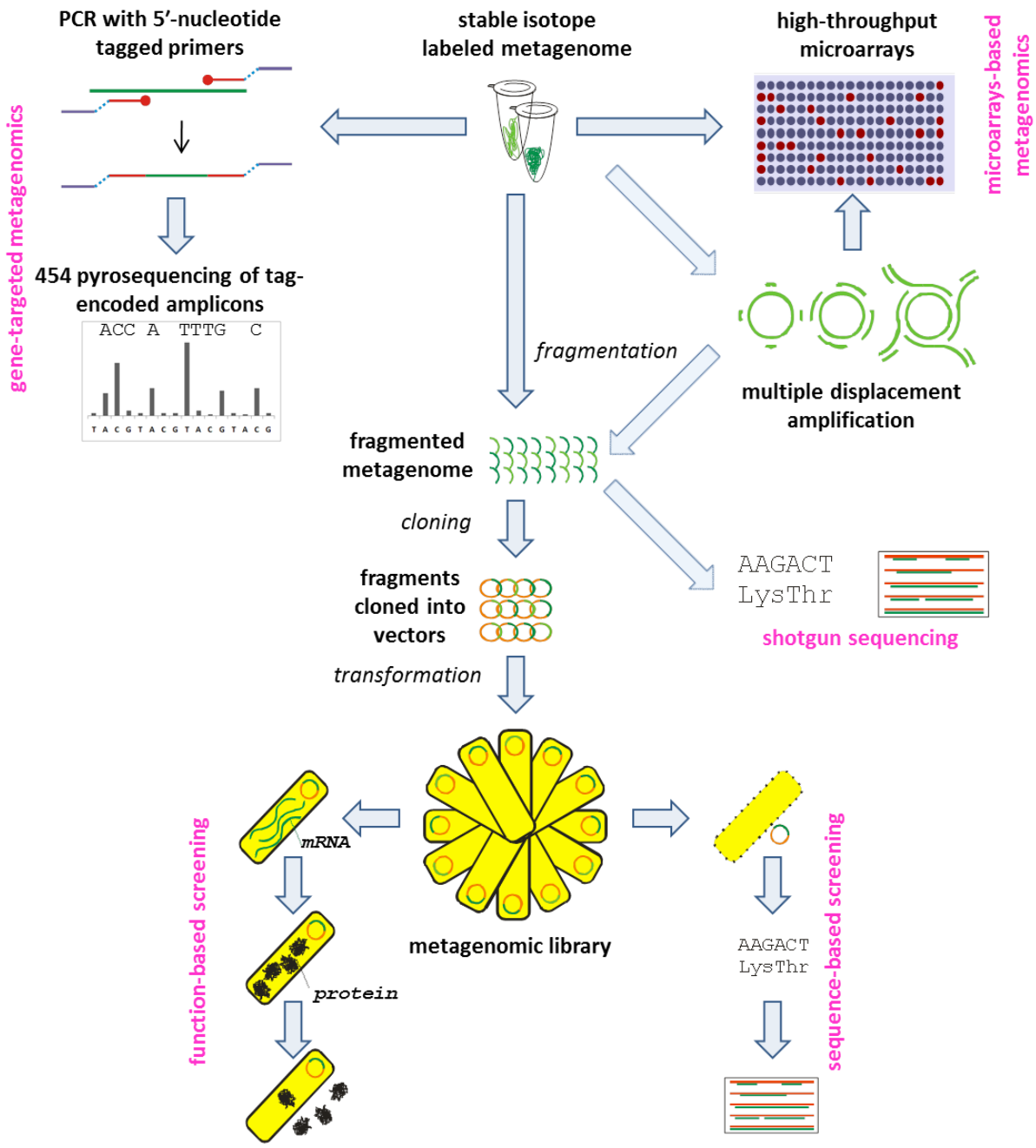

Figure A2-2: Overview of metagenomic approaches that can be used to analyze stable isotope labeled metagenomes.

Table A2-1: Comparison of methodological considerations for DNA-SIP, RNA-SIP, PLFA-SIP, and protein-SIP.

Trait	Comparison of applicability of biomarkers	Explanation
Sensitivity	protein > PLFA > RNA > DNA	DNA-SIP requires 15-20% isotopic enrichment, while protein-SIP only requires 1%. RNA labeling is 6.5 faster than that of DNA.
Incubation time	protein > PLFA > RNA > DNA	Incubation time is directly linked to sensitivity. DNA-SIP is the only technique that requires active cell division requiring the longest incubation periods potentially leading to biases.
Taxonomic resolution	DNA ≈ RNA > protein > PLFA	PLFA-SIP only distinguishes broader taxonomic groups, while DNA or RNA-SIP provides identification to the genus level or below. Databases for protein sequences are more limited than for 16S rRNA genes.
Indication of metabolic activity	protein > RNA > DNA	Proteins are the most explicit indicators of metabolic activity, while DNA only shows the metabolic potential.
Ease of isolation	DNA ≈ PLFA > RNA > protein	Isolation of PLFA and DNA are routinely performed in different matrices, but isolation of RNA and proteins from environmental samples can be very challenging.
Stability	DNA ≈ PLFA > protein > RNA	DNA or PLFA are fairly stable, but proteins may denature, and mRNA is very sensitive to degradation.
Application with ’omics‘	DNA > RNA > protein	The application potential depends upon the developmental stage of the ’omics‘ methods. Currently, metagenomics is the most well-developed followed by metatranscriptomics and metaproteomics, respectively.